

U.S. Department of Commerce
National Oceanic and Atmospheric Administration
National Weather Service
National Centers for Environmental Prediction
5830 University Research Court
College Park, MD 20740-3818

Office Note 508

<https://doi.org/10.25923/pjwn-p075>

Overview of the 2020 Atmospheric Rivers Field Campaign

Stephen J. Lord, Xingren Wu, Vijay Tallapragada

NOAA/NWS/NCEP Environmental Modeling Center

College Park, Maryland

March 2022

E-mail: vijay.tallapragada@noaa.gov

ABSTRACT

The impact of assimilating dropsonde data from the 2020 Atmospheric River (AR) Reconnaissance (ARR) field campaign on operational numerical weather forecasts was assessed. Two experiments were executed for the period from 24 January to 18 March 2020 using the National Centers for Environmental Prediction (NCEP) Global Forecast System version 15 (GFSv15) with a four-dimensional hybrid ensemble-variational (4DEnVar) data assimilation system. The control run (CTRL) used all of the routinely assimilated data and included data from 628 ARR dropsondes, whereas the denial run (DENY) excluded the dropsonde data. Results from 17 Intensive Observing Periods (IOPs) indicate a mixed impact for mean sea-level pressure and geopotential height over the Pacific North American (PNA) region in CTRL compared to DENY. The overall local impact over the U.S. West Coast and Gulf of Alaska for the 17 IOPs is neutral (-0.45%) for Integrated Vapor Transport (IVT), but positive for wind and moisture profiles (0.5-1.0%), with a spectrum of positive and negative impacts for various IOPs. The dropsonde data impact on precipitation forecasts over U.S. West Coast domains is largely positive, especially for day 5 lead time, and appears driven by improved low-level moisture fields. These results suggest that data gaps associated with ARs can be addressed with targeted ARR field campaigns providing vital observations needed for improving U.S. West Coast precipitation forecasts.

1. Introduction

Atmospheric Rivers (ARs), generally described as bands of maximum horizontal water vapor flux typically 200 km in extent and primarily in the lower troposphere, are usually part of a mid-latitude cyclone complex that has high moisture content originating from the central tropical Pacific region. ARs have an important impact on weather over the U.S. West Coast and Canada and are associated with 84% of flood damage in the western USA (Ralph et al. 2020). However, forecasting AR impact on typical watersheds, with an extent 100 km or less, has many difficulties, among which are the comparable horizontal extents of both the AR and the target watershed areas, coastal topographic forcing on the USA and Canadian west coasts, and the inherent challenges of accurate precipitation prediction from operational forecast and data assimilation systems.

Real-time AR Reconnaissance (ARR) Observing Campaigns (OCs, Ralph et al. 2020) have been designed and executed in a similar manner to those initiated for tropical cyclones (e.g., Aberson et al. 2010, Majumdar et al. 2013, Brennan et al. 2015). Moreover, dropsondes deployed in the vicinity of extratropical and tropical cyclones in mid-latitudes have shown positive downstream impact (Schäfler et al. 2018 and Schindler et al. 2020).

In the case of ARs, aircraft are deployed to sample the jet-like features throughout the troposphere and of regions with enhanced moisture emanating primarily from the tropics. Ensemble-based (Torn and Hakim 2008, Elless et al. 2021) and adjoint-based (Demirdjian et al. 2020a, b) sensitivity analyses are used for each AR IOP to direct targeted AR dropsonde observations into regions with potential for improving predictions of U.S. West Coast precipitation. The ensemble sensitivity tools use data from the European Centre for Medium-Range Weather Forecasts (ECMWF), the Canadian Meteorological Center (CMC) and the NCEP Global Ensemble Forecast System (GEFS). In the majority of cases, both the adjoint and ensemble sensitivity analyses indicated forecast sensitivity along the AR core, its edges, and in the warm conveyor belt, all of which provided complementary information where additional observations may improve forecasts (Ralph et al. 2020).

Dropsonde soundings from each aircraft provided temperature, wind, pressure and moisture data that were transmitted to international weather prediction centers in real-time for ingest into

their respective operational data assimilation and forecast systems. The resulting forecasts have the benefit of these data as well as previous soundings (through the cycled data assimilation) and would be expected to be of improved predictive quality for users concerned with ARs and their regional impacts along the U.S. West Coast.

This paper examines the impact of AR dropsonde observations taken during 17 AR IOPs over January-March 2020 (AR2020). The NCEP operational Global Forecast System (GFS), consisting of a forecast model and a cycled data assimilation system, is executed with and without the observations and provides the raw impact data. The impact is measured in terms of reduced error for AR-important weather predictands such as wind, water vapor and precipitation, as well as other standard weather prediction variables, temperature, geopotential height and sea-level pressure.

Section 2 briefly reviews AR OCs prior to 2020 and Section 3 details the methodology of the data impact experiments. In Section 4, the 2020 AR IOPs and fundamental quantities such as Integrated Vapor Transport (IVT, Ralph et al. 2018) and precipitation are described. Section 5 gives the impact results for basic model variables, IVT and precipitation. Appendices A-E provide details on AR2020 IOPs, a list of critical forecasts hours and verification dates for precipitation, a case study of large positive impact resulting from dropsondes deployed during AR2020 IOPs and details on forecast IVT errors and precipitation impacts directly from AR IOPs.

2. Review of prior AR OCs

AR Recon programs have been under development since 2014 to provide aircraft-based sounding observations and other field data, such as deployed buoys, to global operational weather prediction centers and the research community (Ralph et al. 2014). Ralph et al. (2020) summarizes the most recent AR missions from 2016-2019. Participating aircraft have been the U.S. Air Force C130 turboprop aircraft (with a flight ceiling of approximately 450 hPa, range of 3300 km and endurance exceeding 10.5 hours) and the NOAA Gulfstream IV (G-IV) jet (with a flight ceiling of approximately 150 hPa, range of 6600 km and endurance of about 8.75 hours). Each aircraft is equipped with dropsondes that can be released at designated intervals over the flight track to sample temperature, moisture, wind and pressure from flight level to the surface. Ingest of these data into operational systems, with an expected improvement in real-time forecast accuracy, has been evaluated by Stone et al. (2020), Lavers et al. (2018, 2020) and Zheng et al. (2021).

Hamill et al. (2013) studied the impact from the assimilation of targeted observations from the 2011 Winter Storms Reconnaissance (WSR) Program with parallel cycles of ECMWF's data assimilation and deterministic forecasts by including or excluding the targeted observations with the rest of the regularly assimilated data. However, they found that the 2011 WSR results do not support the hypothesis that differences between forecasts with and without the assimilated dropsondes are statistically (significantly) improved in the localized verification region. They noted that there may be several reasons for the lack of impact in their study, including improvement in the observing systems and the data assimilation and forecast systems, and the incomplete dropsonde sampling of the initial sensitive area due to limitations on deploying aircraft range and other flight restrictions. In addition, they reported that the Ensemble Transform Kalman Filter (ETKF) targeting technique used for the WSR experiment was imperfect and also inconsistent with the operational data assimilation scheme used in their study. Nevertheless, the work reported here adopts a similar ETKF methodology, but with the NCEP GFS, as described below.

3. Methodology

a. Data assimilation and forecast model

The experimental methodology is that of a standard NWP data impact experiment. One cycled data assimilation and forecast run is executed with the additional AR observations (the CTRL) and a second run (DENY) is made without these observations. All other aspects of the analysis and forecast system and generated data are the same. A set of diagnostic error post processing software is applied to each run and differences in verification statistics are generated and evaluated. The forecast model and data assimilation system used to generate the CTRL and DENY impact experiments is the GFS version 15 (GFSv15), which was an upgrade over the version used in NCEP's operations during most of the AR2020 OC.

The GFSv15 model (Yang and Tallapragada 2018) has a horizontal resolution of 13 km, and 64 levels in the vertical extending up to 0.2 hPa. The GFDL finite-volume cubed-sphere (FV3) dynamical core (Lin and Rood 1997, Lin 2004, Putman and Lin 2007, Harris and Lin 2013, Harris et al. 2020a, b) is the basis of GFSv15, and a suite of physical parameterizations comprise the GFS model. The GFSv15 upgraded physical parameterization package includes the replacement of Zhao-Carr microphysics with the more advanced GFDL microphysics (Zhou et al. 2019), an updated parameterization of ozone photochemistry with additional production and loss terms (McCormack et al. 2006), a newly introduced parameterization of middle atmospheric water vapor photochemistry (McCormack et al. 2008), a revised bare soil evaporation scheme to reduce a dry and warm bias, and a modified convection scheme to reduce excessive cloud top cooling.

The Global Data Assimilation System (GDAS) is a 4D-Hybrid Ensemble-Variational Data Assimilation System (Kleist and Ide 2015b). The ensemble system has 80 members at a resolution of 25 km. The Hybrid algorithm combines uncertainty estimates from the ensemble and a climatological, spatially varying but fixed in time, variance.

The operational GDAS observations include hyperspectral polar-orbiting and geostationary sounder and imager radiances, radiosonde soundings, GPS radio-occultation soundings, atmospheric motion vectors derived from geostationary imagers, buoy and ship observations and land-based surface observations (Kleist et al. 2009, Kleist and Ide, 2015a, b). The GDAS is cycled 4 times daily for data centered at 00, 06, 12 and 18 UTC. The observation window is plus/minus 3 hours surrounding each cycle time. A 9 hour forecast from the previous cycle is used as a background field in the hybrid assimilation scheme and a variational quality control algorithm is used to down-weight suspicious data.

The experimental period began at 00 UTC 24 January 2020 and continued until 00 UTC 18 March, covering 55 days, of which 17 days have IOP dropsonde observations. Since the final IOP was at 00 UTC 11 March, forecasts from both CTRL and DENY experiments out to 168 h (7 days) can be verified against their own analyses. The vast majority of dropsonde launches occurred surrounding the 00 UTC observation window, so that sondes launched during the 06 and 18 UTC observation windows will not be explicitly referenced (Table A1).

b. Precipitation Observations

Precipitation observations are critical for verifying AR impact. Over land, the standard operational Climatologically Calibrated Precipitation Analysis (CCPA) Version 4 is used (Hou et al. 2014). The CCPA is a regression-based merging of the Climate Prediction Center Unified Global Daily Gauge Analysis and the Environmental Modeling Center's Stage IV multi-sensor precipitation product (Lin and Mitchell, 2005). The CCPA product is available twice daily (00 and 12 UTC) and is a 24- hour accumulated quantity over the CONUS on a 0.125 degree latitude-longitude (~10 km at 40 N) grid. Over the ocean, the Climate Prediction Center Morphing technique (CMORPH, Joyce et al. 2004), a satellite-based product, is used. CMORPH uses Level 2 precipitation rate retrievals from passive microwave instruments aboard low-earth orbiting

satellite platforms and infrared brightness temperatures from geostationary platforms and infrared brightness temperatures from geostationary platforms

(https://www1.ncdc.noaa.gov/pub/data/sds/cdr/CDRs/Precipitation-CMORPH/AlgorithmDescription_01B-23.pdf). CMORPH is defined quasi-globally (60S to 60N) on a 0.25 degree latitude-longitude grid every 30 minutes. The CMORPH data is accumulated over 24 hours at 00 and 12 UTC to match the CCPA product. No attempt is made to reconcile the CCPA and CMORPH products at the coastline; the CMORPH product can be biased and is definitely not as reliable over land as CCPA. Precipitation dates are identified by the end of their accumulation, viz. a map labelled “12 UTC 27 January” shows the 24 h accumulation (mm) ending on that date.

c. Experimental design

Two GFS experiments were executed, each consisting of a 6 hourly GDAS cycle over the experimental period and forecasts out to 168 h generated from initial conditions at 00 and 12 UTC daily through 00 UTC 11 March 2020, the last IOP date. The control experiment (“CTRL”) assimilated all dropsonde observations received operationally by NCEP and the second experiment (“DENY”) did not assimilate any dropsondes. Output data consisted of global analyses and forecast fields on a 1x1 latitude/longitude grid except for precipitation, which is on the model Gaussian grid (~13 km), and all assimilated observations. As such, CTRL output is very similar to the operational GFS output, with much the same error patterns at all scales. By comparing error patterns and statistics from CTRL and DENY experiments, one can assess the impact of the AR OC data for 2020. Similar data from the operational ECMWF (“ECMO”) system was also used to provide independent verifying analyses for determining forecast error and to provide additional insight into the CTRL and DENY experiments.

The 17 IOP cases represent approximately 30 percent of the total cases (55) over the experimental period. Thus, statistical mean values over all verifications can dilute the impact of the IOP cases alone. The approach taken here is to examine mean statistics over all 55 initial conditions but to also examine impacts over the 17 IOP cases separately. Similarly, for geographical extent, the observation impacts over North America are likely to be small, having been diluted by a large areal extent compared to the area covered by and impacted by the observations. Again, the approach taken here is to briefly address impact over North America, but to focus on the impact over a much smaller domain, the NE Pacific and U. S. West Coast, which is also most consistent with the goals of the AR Program.

The IVT is a fundamental measurable for ARs. It is defined as:

$$IVT = Mag \left\{ \left[\int_{sfc}^{225} (u * q) dp / g \right], \left[\int_{sfc}^{225} (v * q) dp / g \right] \right\} \quad (1)$$

where sfc is the surface pressure, 225 hPa is the upper integration limit, Mag is the magnitude operator, u, v are respectively the horizontal zonal and meridional velocities (m/s), q is the specific humidity (g/g) and g is the acceleration of gravity.

CTRL and DENY were evaluated visually by their error patterns with respect to ECMO verifying analyses at 00 and 12 UTC throughout the experimental period and statistically using a subset of NCEP’s standard statistical scores and mean absolute error (MAE) statistics for IVT and precipitation. Other selected model variables were also evaluated such as mean sea-level pressure (abbreviated as PMSL), geopotential height at 500 hPa (Z500), wind speed (WSPD) at mandatory pressure levels from 200-925 hPa, and specific humidity (SPCH) at 700, 850 and 925 hPa. Table 1 summarizes the variety of statistical results calculated over appropriate verifying domains (see also Fig. 1) and presented in this paper.

Domain Number	Time Domain	Spatial Domain & Projection	Analyzed Variables (abbreviation)	Levels (hPa)	Description
1	00 UTC 24 January 2020 – 00 UTC 18 March 2020	Large domain (180-320 E, 20-75 N, lat-lon projection)	Geopotential height (Z) Wind speed (WSPD)	500 500	TMS
2	00 UTC 24 January 2020 – 00 UTC 18 March 2020	Large domain (163-277 E, 14-75 N, polar stereographic projection)	Geopotential height (Z) Mean sea-level pressure (PMSL)	500 Surface	TMS
1	00 UTC 29 January	Domain 1	Geopotential height (Z)	500	CS
3	17 IOP dates	Small domain (200-245 E, 33-65 N, lat-lon projection)	Geopotential height (Z) Mean sea-level pressure (PMSL) Wind speed (WSPD) Specific humidity (SPCH)	200 500 700 850 Surface 200 300 400 500 700 850 925 700 850 925	FS for IOP cases
4	As in Domain 3	Small domain (180-250 E, 18-60 N, lat-lon projection) traversed by all IVT maxima	Integrated vapor transport (IVT)	250	FS for all IOP cases
5	00 and 12 UTC for all dates from 25 January to 16 March	West Coast (WEST) regional domain (233-253 E, 28-49.5 N, lat-lon projection)	Precipitation	Surface	FS and QCD
6	As in Domain 5 and 17 IOP dates and critical forecasts (Table B1) over the period 00 UTC 25 January to 12 UTC 18 March	Pacific Northwest and Northern California (PNNC) regional domain (235-243 E, 36-49.5 N, lat-lon projection)	Precipitation	Surface	FS and QCD
7	As in domain 6	Southern California, Arizona, New Mexico (SCAN) regional domain (238-265 E, 28-39 N, lat-lon projection)	Precipitation	Surface	FS and QCD
8	IOP-13 (00 UTC 7 March)	Baja California (230-300 E, 15-35 N, lat-lon projection)	IVT, precipitation	Surface	QCD

Table 1. Summary of the time and space domains used in calculating verification statistics. Analyzed variables (and their abbreviations), vertical levels and a short description are also listed. Time-mean statistics

derived from large domains (1) and (2) are comparable, even though the domains are not identical. Appendix A contains the valid dates for each IOP. Descriptions are abbreviated (TMS: time-mean statistics for the OC; CS: case study; FS: forecast statistics; and QCD: qualitative case description). Spatial domains are shown in Fig. 1.

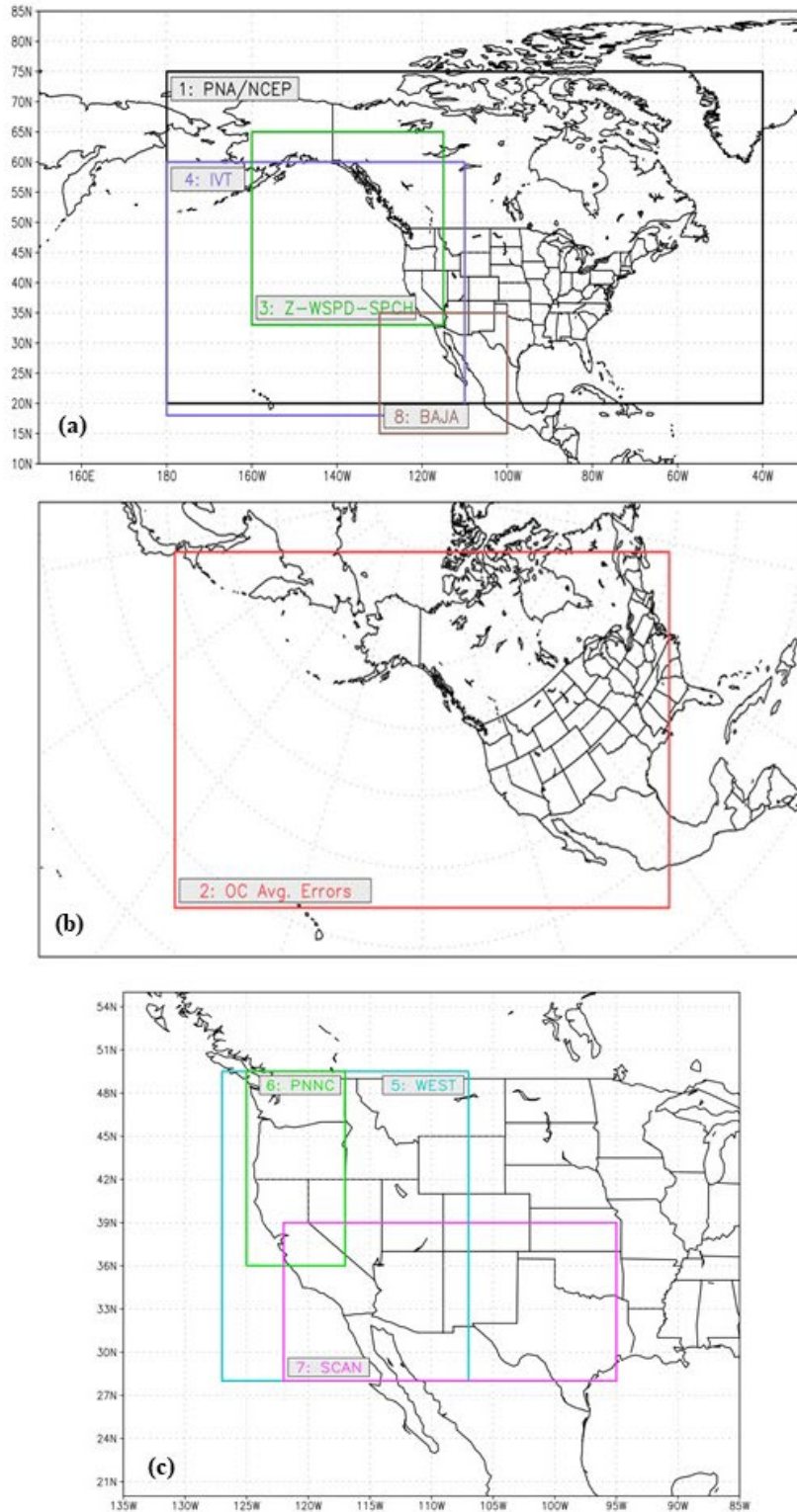


Fig. 1. Spatial domains 1, 3, 4 and 8 (a) for the NCEP statistical, model variable (Z, WSPD, SPCH) and IVT verifications respectively (see Table 1), domain 2 (b) for the large-scale verification for PMSL and Z over the AR OC and domains 5-7 (c) for precipitation verification.

d. Estimating and evaluating data impacts

Since dropsonde data are ingested into a continuously cycled data assimilation system, typically with 4 cycles/day, impact can occur at the cycle with assimilated AR observations due to an improved initial atmospheric state at the assimilation time, or it may occur at a later data assimilation cycle due to spread of observed information throughout the model domain as the continuous cycling operates over time, even though there are no new AR observations assimilated. In the first case, one would expect improvement in the short-range forecasts (12-72 h). In the latter case, a forecast improvement may occur at any time from a cycle when dropsondes are not assimilated. Due to the continuous nature of a cycled data assimilation system, it is likely that both impacts may be realized, but “direct impact” from dropsonde data at the initialization (analysis) time is more common than “indirect impact” from previously assimilated data. Separating direct and indirect impacts is extremely difficult. While direct impacts are more likely to be positive, indirect impacts from remote areas may be negative and may overwhelm direct impacts.

The magnitude of data impact is, however, extremely difficult to predict. Due to the possibility of fast-growing, non-linear errors in the forecast model, any perturbation from dropsonde data can result in a large (positive or negative) impact, even though many/most impacts are very small. Sampling the atmosphere over a limited geographical area with time-separated observations based on phenomenological considerations has substantial merit but does not ensure success due to indirect impacts as noted above. Nevertheless, while the question of observation impact is somewhat murky, for the purposes of this study, we will adopt a simplified strategy by defining direct and indirect impact as above.

4. The AR2020 Observing Campaign and IOPs

The AR2020 OC took place over the Northeast Pacific and Gulf of Alaska from 24 January to 11 March 2020. Seventeen IOPs were conducted using the NOAA Gulfstream IV jet (G-IV) and two C-130 turboprop aircraft from the Air Force 53rd Weather Reconnaissance Squadron’s (AFRES’s) Hurricane Hunter group as described above. Dropsondes were deployed during each IOP and the data were transmitted in real-time to operational weather prediction centers, including NCEP. Dropsonde temperature and moisture sensor accuracy is approximately that of radiosonde data taken routinely and world-wide over landed areas and scattered islands by operational national weather services; dropsonde wind errors may be slightly larger than for radiosondes. All IOPs begin on the 00 UTC cycle in year 2020 so that dates are identified by their calendar day and month (e.g., 5 February). Additional details on the 17 AR2020 IOPs are in Appendix A.

The IOPs were initiated based on satellite-based evidence of an existing AR and real-time, operational forecasts predicting their future impacts over the western USA. To make precipitation verification most relevant to the impacted geographical area, three special domains were used for precipitation (Table 1). The first domain covers the Pacific Northwest (Washington and Oregon) and northern California (PNNC), the second domain covers southern California, Arizona and New Mexico (SCAN), and the third domain covers the West Coast region and encompasses both PNNC and SCAN domains. The observed CCPA domain mean and maximum 24 h accumulated precipitation (mm) within the WEST, PNNC and SCAN domains (Fig. 1 and Table 1) show extreme values within 1-2 days following (or during) each IOP. Maximum area-average precipitation for the WEST domain in this time window covering all IOPs is 2-4.5 mm and up to 10 mm for the PNNC domain and 7 mm for the SCAN domain. Maximum precipitation accumulations in the WEST domain exceed 200 mm and occur in the PNNC domain during both early and late IOPs (Fig. 2e) and in the SCAN domain during late February IOPs and March IOPs (Fig. 2f). Since the WEST domain covers both PNNC and SCAN domains, comparison of each

sub-domain with the WEST domain shows that they cover all major precipitation events over the OC. A short summary of the synoptic conditions, the AR IVT horizontal distribution and accompanying precipitation is presented below for each IOP, beginning with IOP-1 at 00 UTC 24 January. The critical forecast ranges (h) for each IOP, which cover the dates of possible high impact precipitation events subsequent to each IOP, are tabulated in Appendix B.

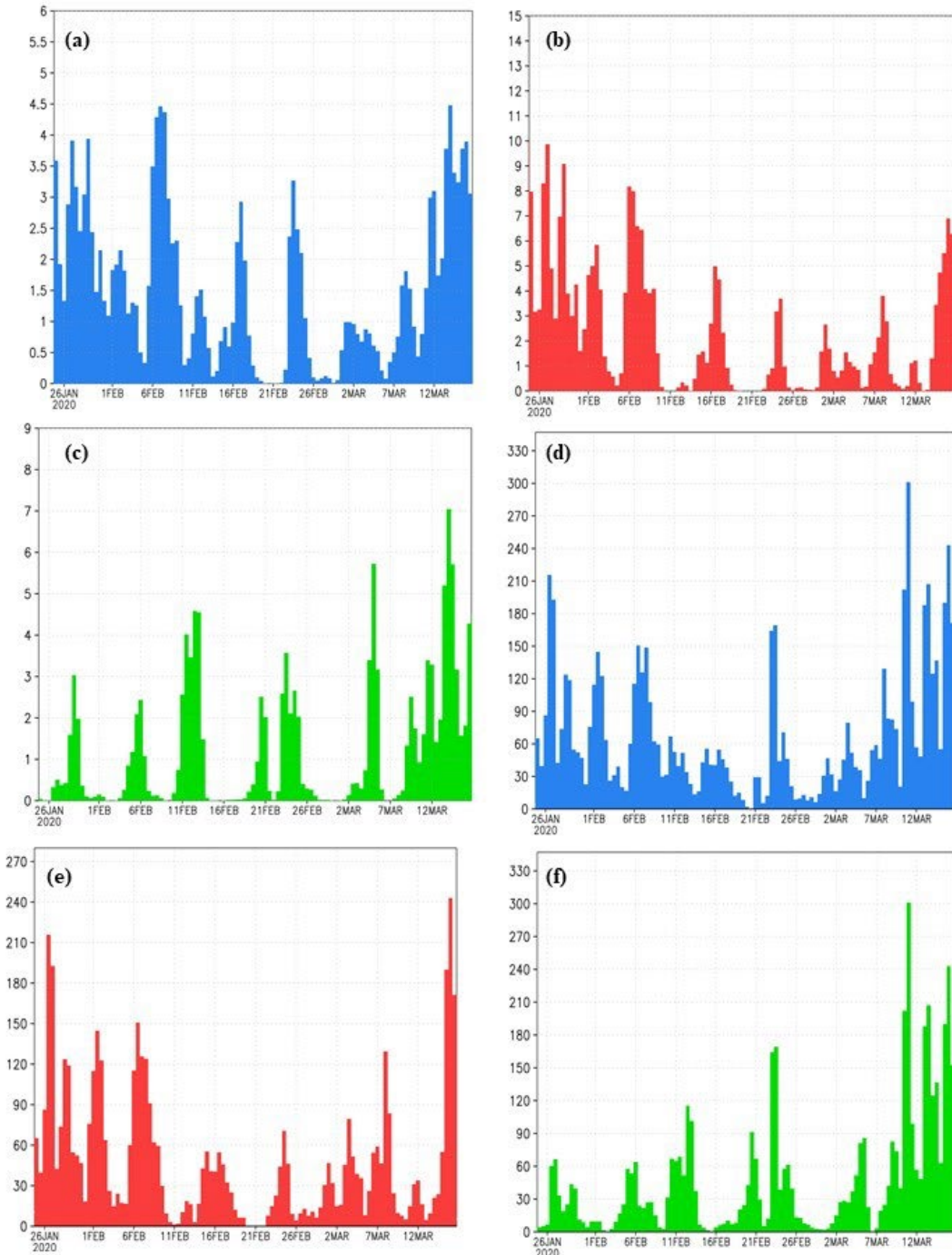


Fig. 2. Area-mean, 24 h accumulated, observed CCPA precipitation (mm) at 00 and 12 UTC over the WEST (a), PNNC (b) and SCAN (c) domains (Table 1) and domain-maximum 24 h accumulated precipitation over the same three domains (d, e, f) from 00 UTC 25 January to 12 UTC 16 March.

a. IOP-1: 24 January

At the time of the first AR2020 IOP, a precipitation event is already impacting the Oregon-Washington coast with an associated IVT maximum extending offshore and southwest flow along the coast (Fig. 3). A major observed AR feature is centered at 35N, 148W between a low-pressure system to the northwest and a high centered at 23N, 135W. Two AFRES flights sampled this AR, which continues to propagate NE before the leading edge influences the Oregon, Washington, and northern California area at 00 UTC 26 January (Fig. 4) and produces peak precipitation at 12 UTC 26 January (Figs. 2, 5).

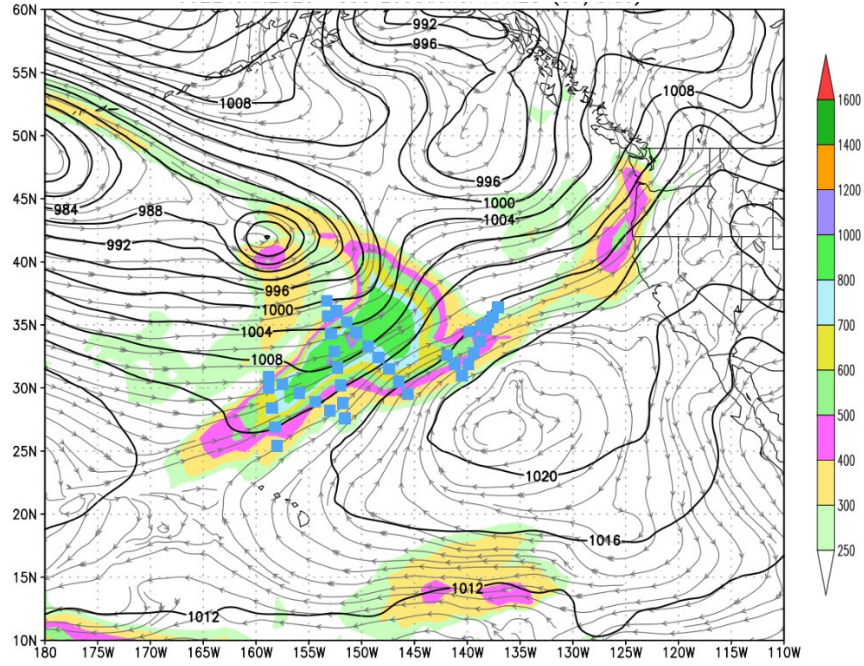


Fig. 3. Magnitude of the CTRL vertically-integrated specific humidity flux (kg/m/s) for 00 UTC 24 January 2020. Contours of mean sea-level pressure (hPa) and 850 hPa streamlines are included. The sounding locations for 37 dropsondes from two AFRES aircraft for the 00 UTC data assimilation cycle are shown in blue.

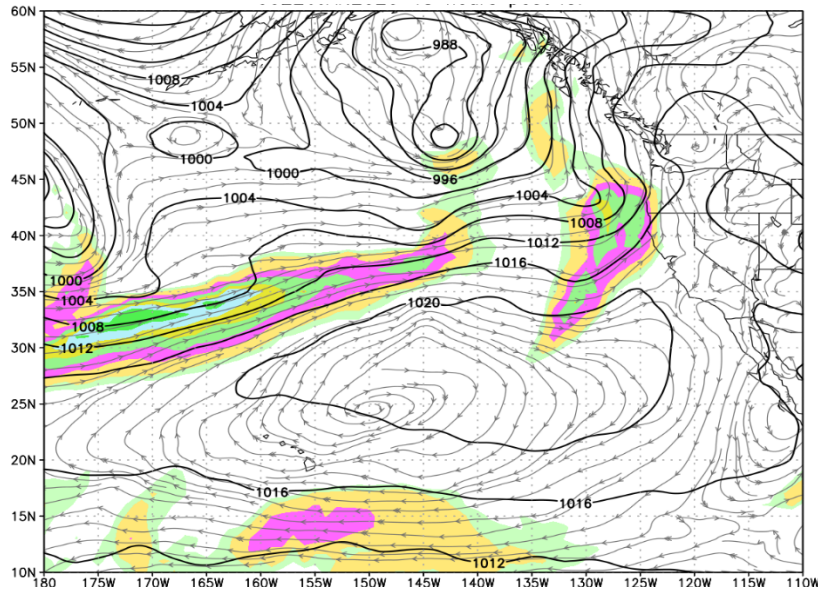


Fig. 4. As in Fig. 3, except valid at 00 UTC 26 January 2020, 48 h after IOP #1.

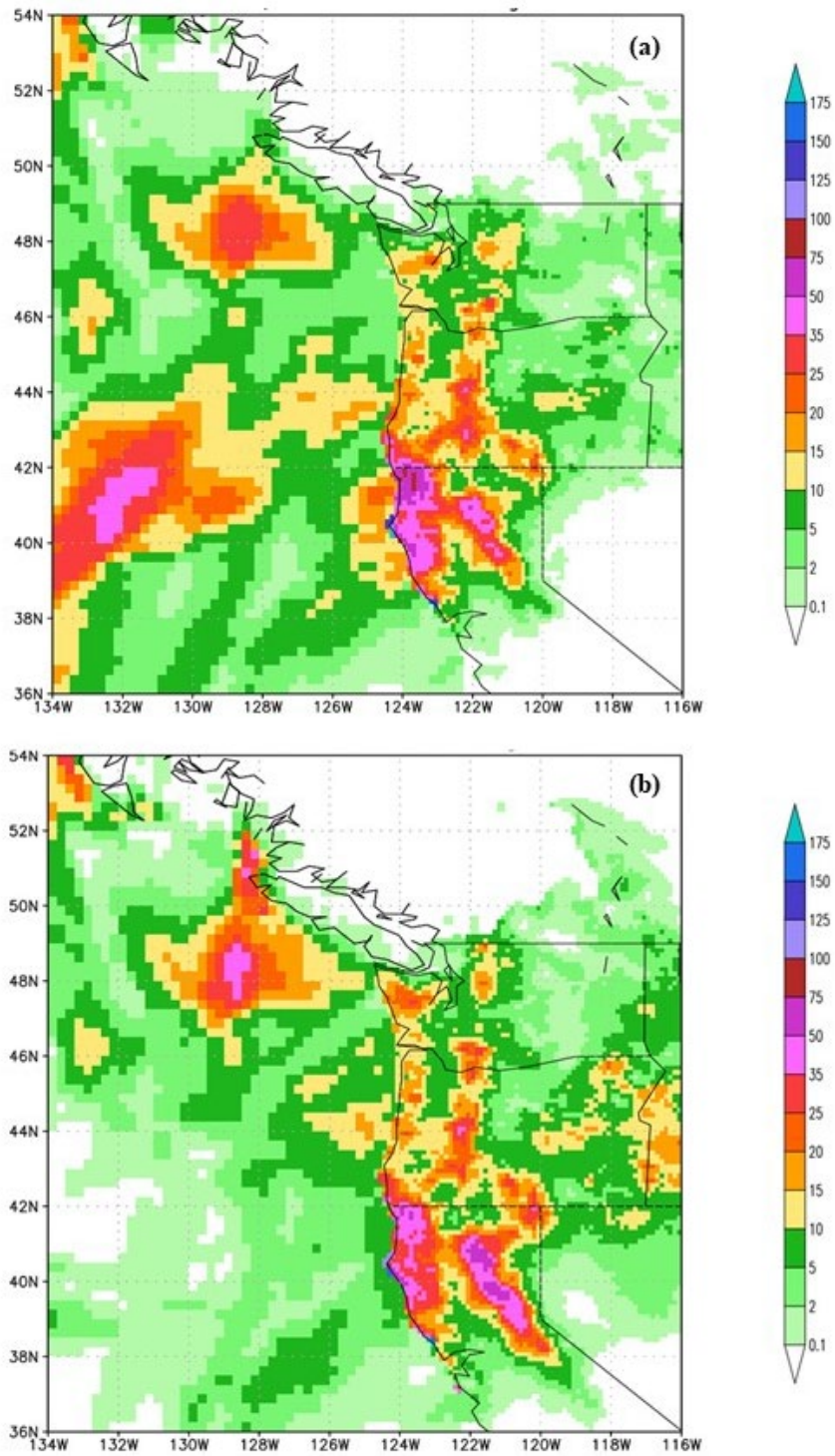


Fig. 5. 24 hour accumulated observed precipitation (mm) ending at 12 UTC 26 January (a) and 00 UTC 27 January (b). Over land is observed CCPA and over ocean is CMORPH satellite-estimated precipitation.

b. IOPs 2, 3: 29, 31 January

A strong low pressure system is propagating across the Gulf of Alaska, centered at 52N, 148 W at 00 UTC 29 January (Fig. 6). An IVT maximum is SSE of the cyclone center and begins to

impact the British Columbia (BC) coast by 12 UTC. A single AFRES mission dropped 21 sondes covering the rear of this system. While land-based precipitation is not available over BC, the CMORPH shows major impact offshore from 12 UTC 29 January until 00 UTC 30 January (Fig. 7a, b). Westerly winds across the Oregon-Washington coastline are consistent with significant precipitation from UTC 28 January 00 UTC 29 January. At 00 UTC 31 January, AFRES soundings continue to sample the area of maximum IVT (Fig. 8); the associated precipitation event lasts from 12 UTC 31 January through 00 UTC 2 February (Fig. 9a-d).

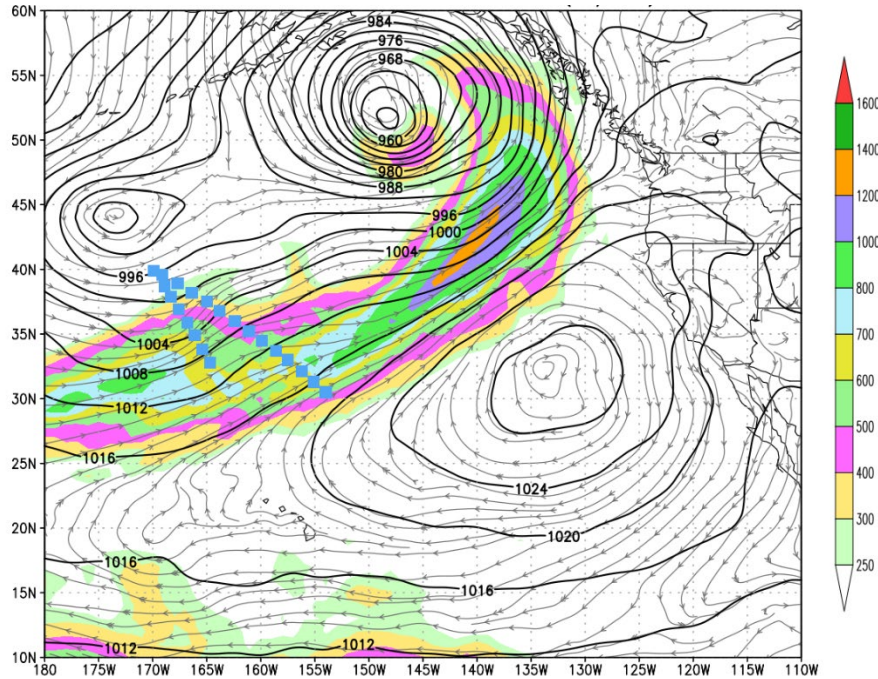


Fig. 6. As in Fig. 3, except for 00 UTC 29 January 2020. Sounding locations from a single AFRES aircraft (21) are shown.

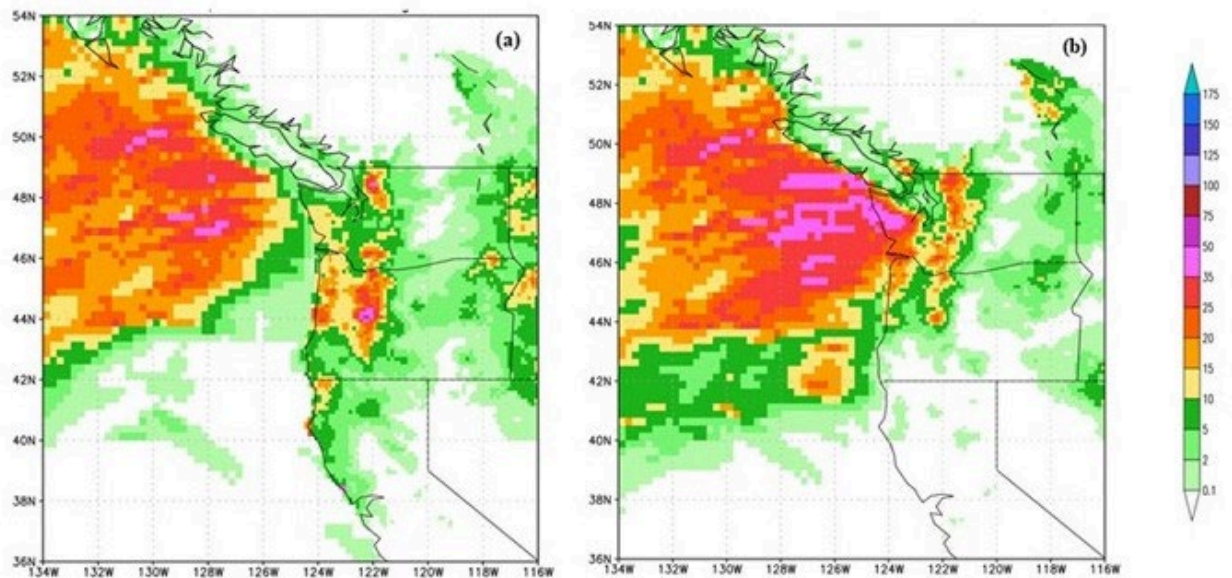


Fig. 7. As in Fig. 5, except for valid date 12 UTC 29 January (a) and 00 UTC 30 January (b).

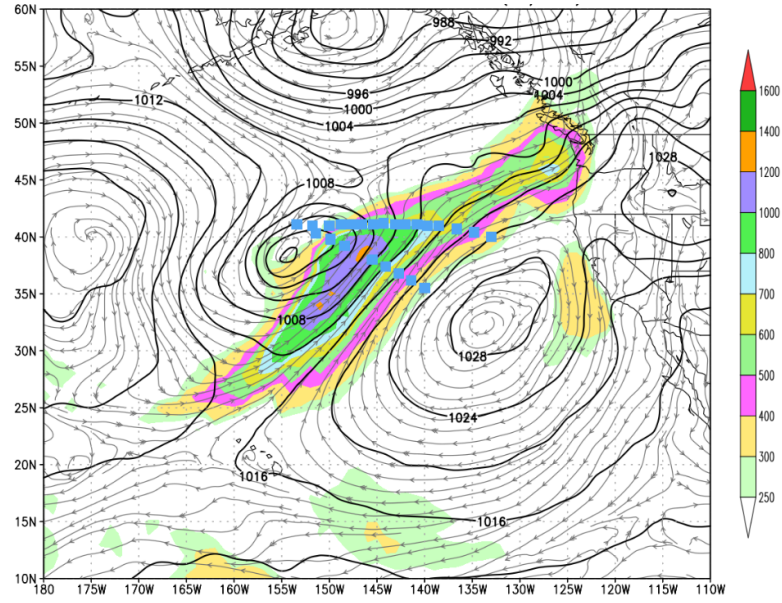


Fig. 8. As in Fig. 3, except for 00 UTC 31 January 2020. Sounding locations from a single AFRES aircraft (24) are shown.

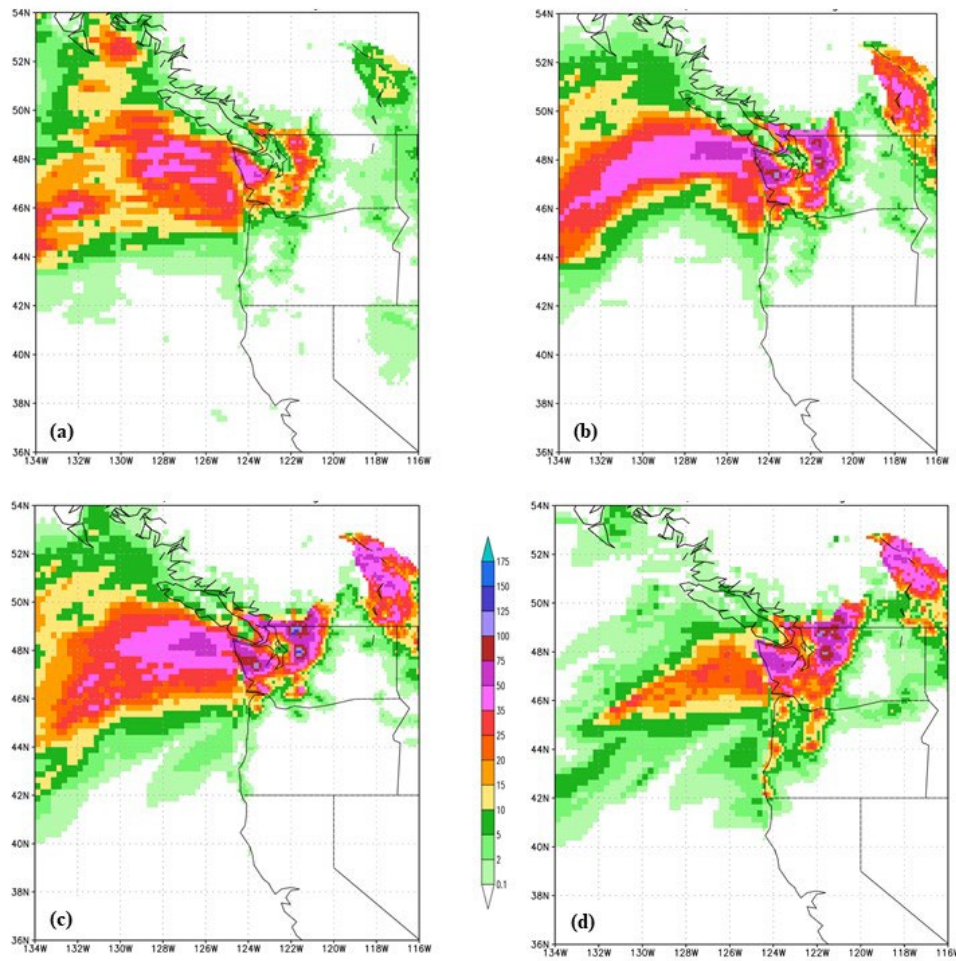


Fig. 9. As in Fig. 5, except for valid date 12 UTC 31 January (a), 00 UTC 1 February (b), 12 UTC 1 February (c) and 00 UTC 2 February (d).

c. IOPs 4-6: 4-6 February

In much the same scenario as 29-31 January, the focus IVT maximum approaches western BC at 00 UTC 4 February (Fig. 10a), but rotates clockwise and landfalls in Washington state by 00 UTC 6 February (Figs. 10b, c). Both AFRES (40) and G-IV (30) sorties contribute soundings for 00 UTC 4 February. The G-IV provides 30 soundings on 00 UTC 5 February and both AFRES (29 from 2 aircraft) and G-IV (30) contribute for 00 UTC 6 February. CMORPH accumulated precipitation estimates show major impact immediately offshore by 00 UTC 5 February and through 00 UTC 6 February (Figs. 11a, b, c) and CCPA shows accumulated impacts over land beginning 12 UTC 5 February (not shown) and at 00 UTC through 9 February (Figs. 12a-c). As the developing cyclone approaches (Fig. 13), westerly to northwesterly winds across the Oregon-Washington coastline are consistent with significant precipitation over Washington State during this period.

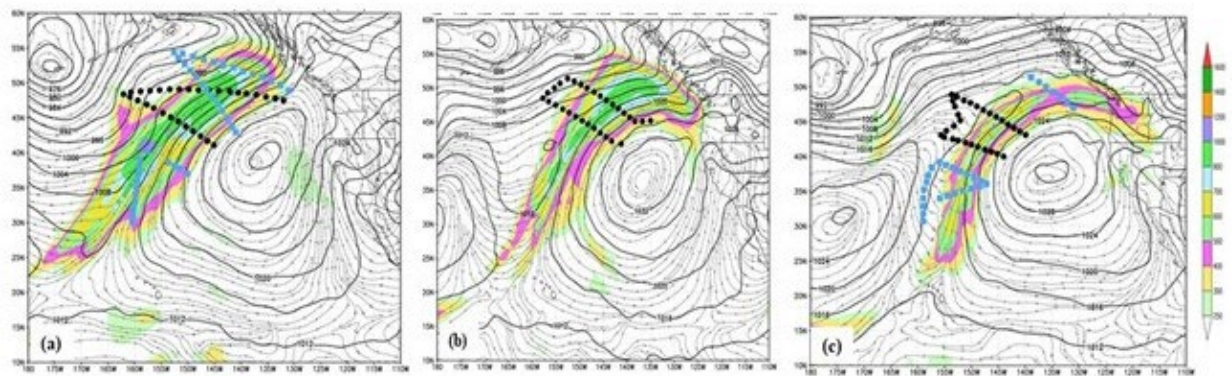


Fig. 10. As in Fig. 3, except for 00 UTC on 4 February (a), 5 February (b) and 6 February (c). Sounding locations from 2 AFRES aircraft (40, blue) and the NOAA G-IV (30, black) on 4 February, the NOAA G-IV (30) only on 5 February, and 2 AFRES aircraft (29) and the NOAA G-IV (30) on 6 February are shown.

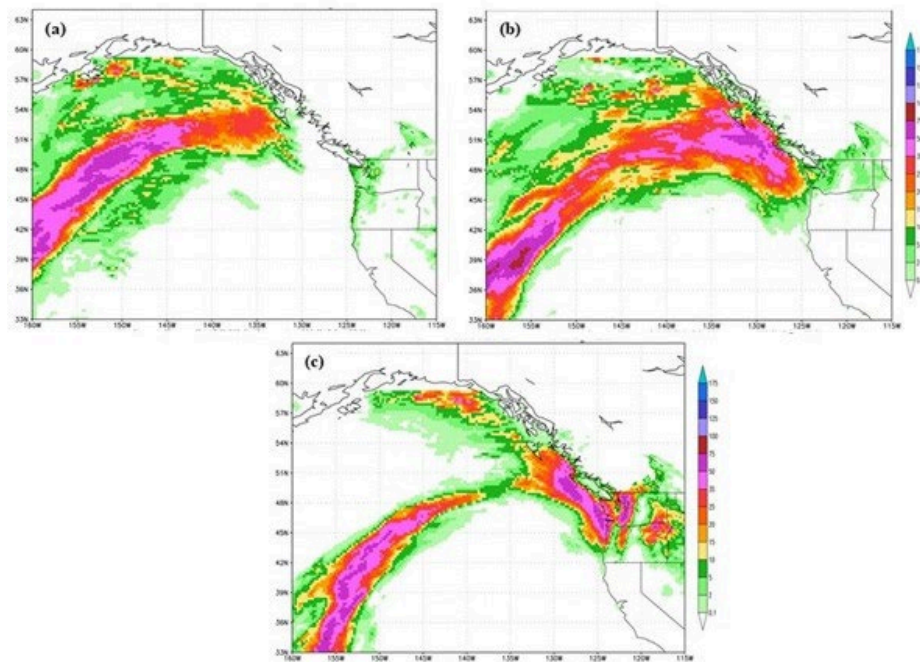


Fig. 11. As in Fig. 5, except for an expanded domain and valid dates 00 UTC on 4 February (a), 5 February (b) and 6 February (c).

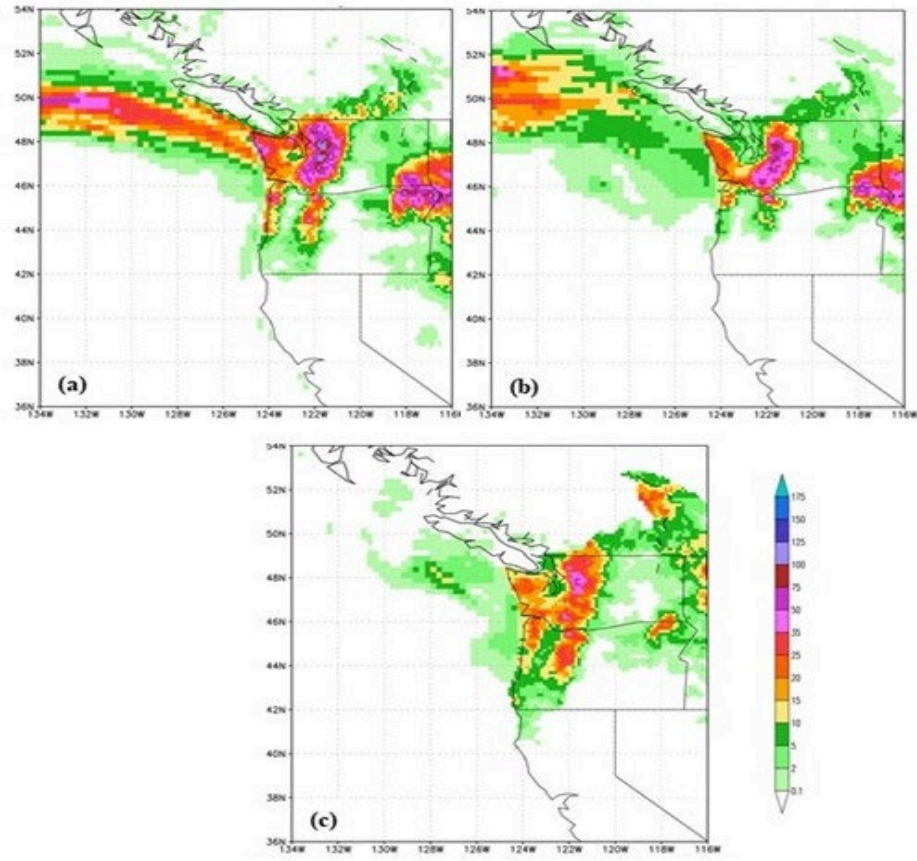


Fig. 12. As in Fig. 5, except for valid dates 00 UTC on 7 February (a), 8 February (b) and 9 February (c).

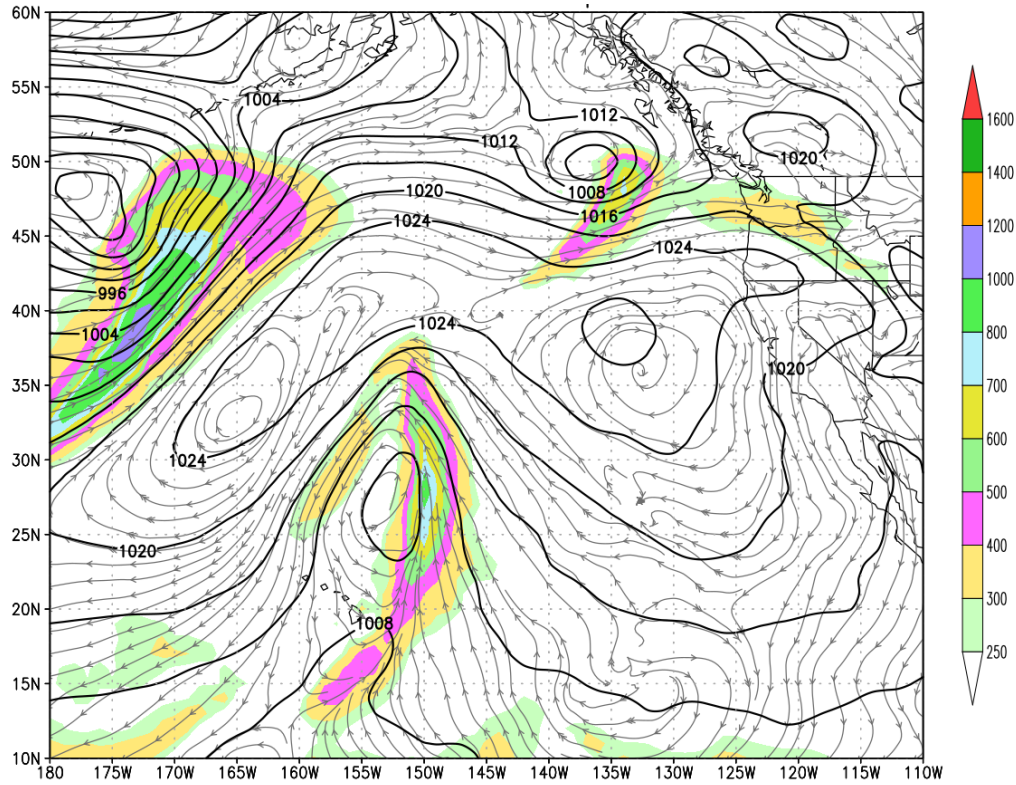


Fig. 13. As in Fig. 3, except for 12 UTC 7 February 2020, 36 h after IOP #6.

d. IOPs 7-9: 14-16 February

The leading edge of the IVT maximum is sampled by AFRES aircraft on 00 UTC 14 February (Fig. 14a). The IVT maximum rotates around the high pressure system to the south to eventually make landfall over the Washington-Oregon area (Figs. 14 b, c). AFRES (48) and G-IV (30) soundings provide a good sampling of the IVT feature over 14-16 February. Precipitation associated with this system begins to impact Washington State on 00 UTC 16 February and then Oregon through 00 UTC 17 February (Fig. 15, a-c). This system is notably less impactful than those discussed in sub-sections a-c above, consistent with the smaller amplitude of the IVT maximum.

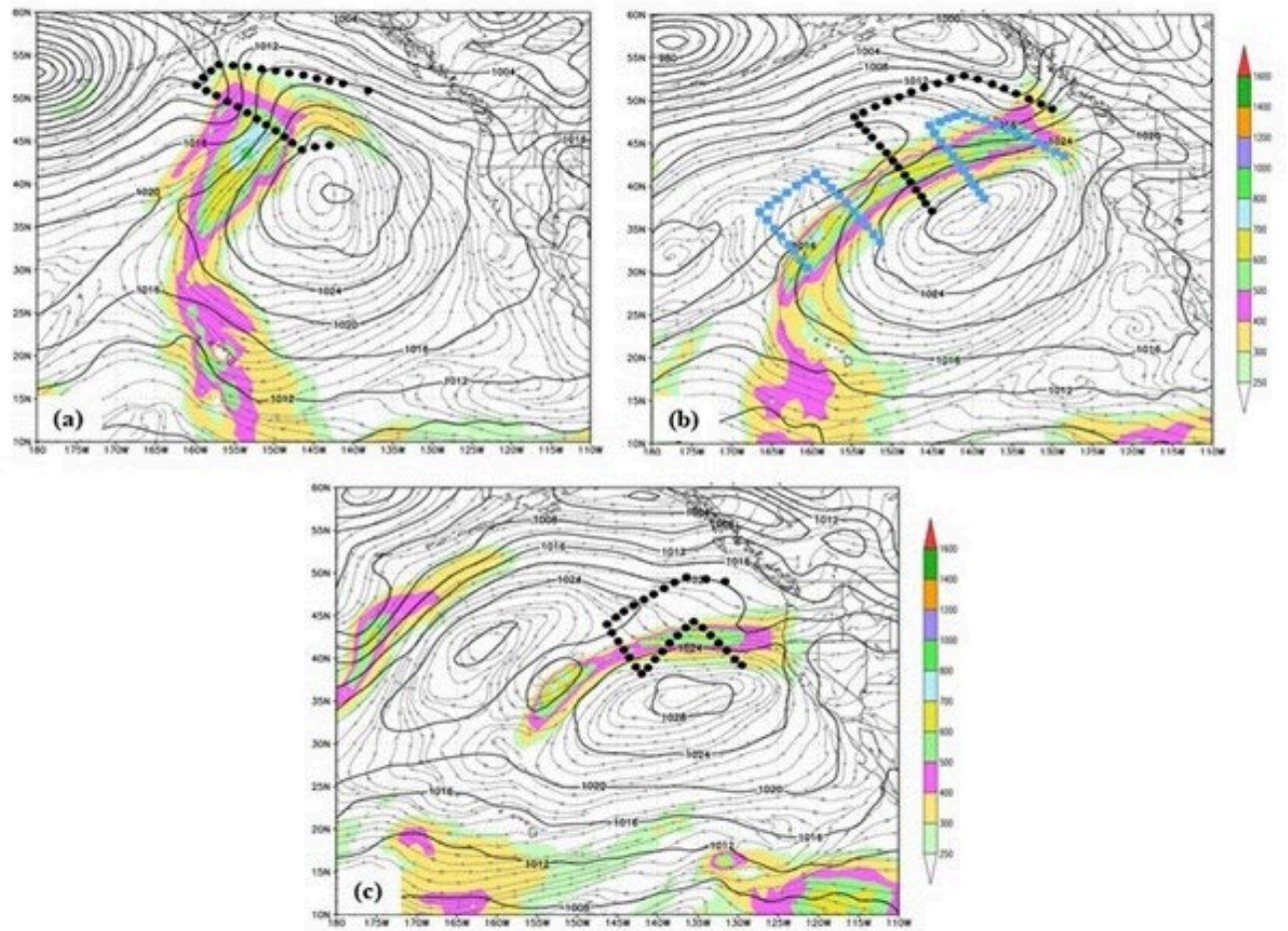


Fig. 14. As in Fig. 3, except for 00 UTC on 14 February (a), 15 February (b) and 16 February (c). Sounding locations from a single AFRES aircraft (27), 2 AFRES aircraft (48) and the NOAA G-IV (30) and the NOAA G-IV (30), respectively, are shown.

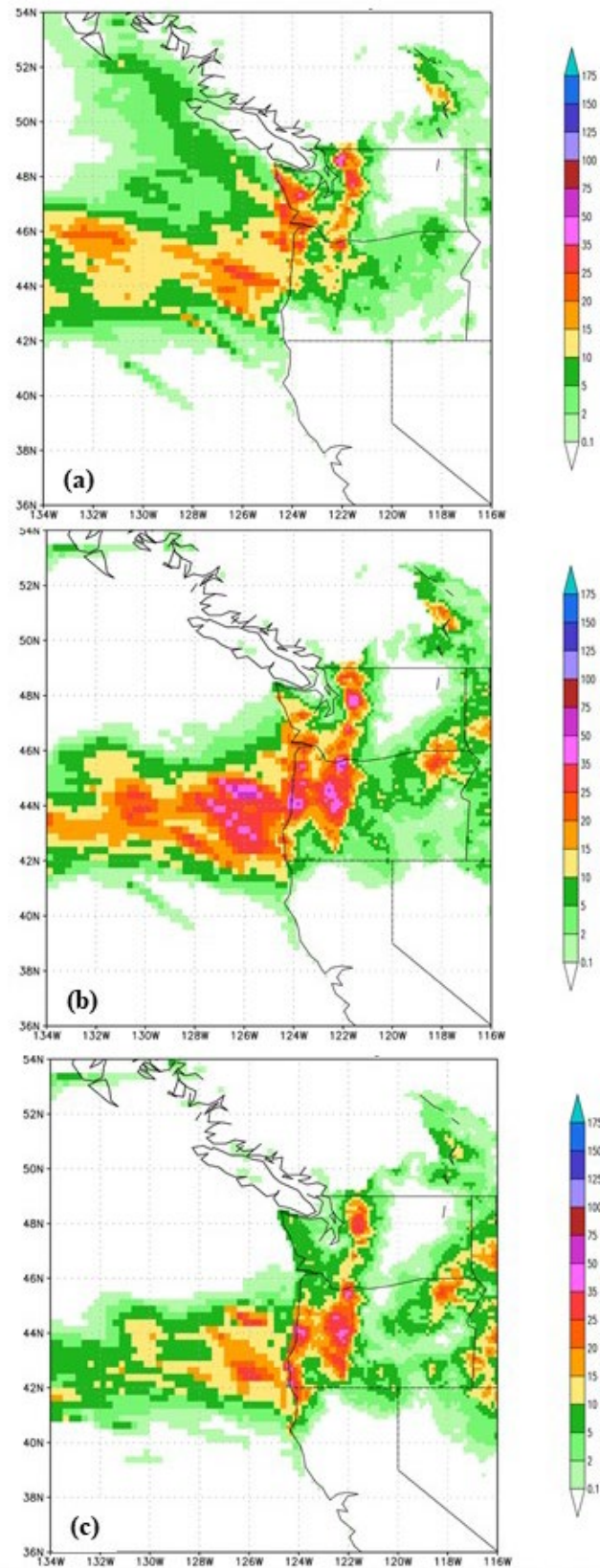


Fig. 15. As in Fig. 5, except for valid dates 00 UTC 16 February (a), 12 UTC 16 February (b) and 00 UTC 17 February (c).

e. IOP-10: 21 February

IOP-10 on 21 February is characterized by an IVT maximum approaching the BC coast, associated with a decaying cyclone in the NW Gulf of Alaska (not discussed further), a later (00 UTC 23 February) developing cyclone offshore of Vancouver Island (Fig. 16a) and an incipient event off Southern California, which was sampled by a single ARFES sortie (Fig. 16b). The tropical origin of this latter event is clearly apparent as it rotates cyclonically on the northwest side of a high pressure system before making landfall over Baja California at 12 UTC 21 February. The precipitation impact is over southern Arizona from 12 UTC 22 February to 12 UTC 23 February (Fig. 17a, b).

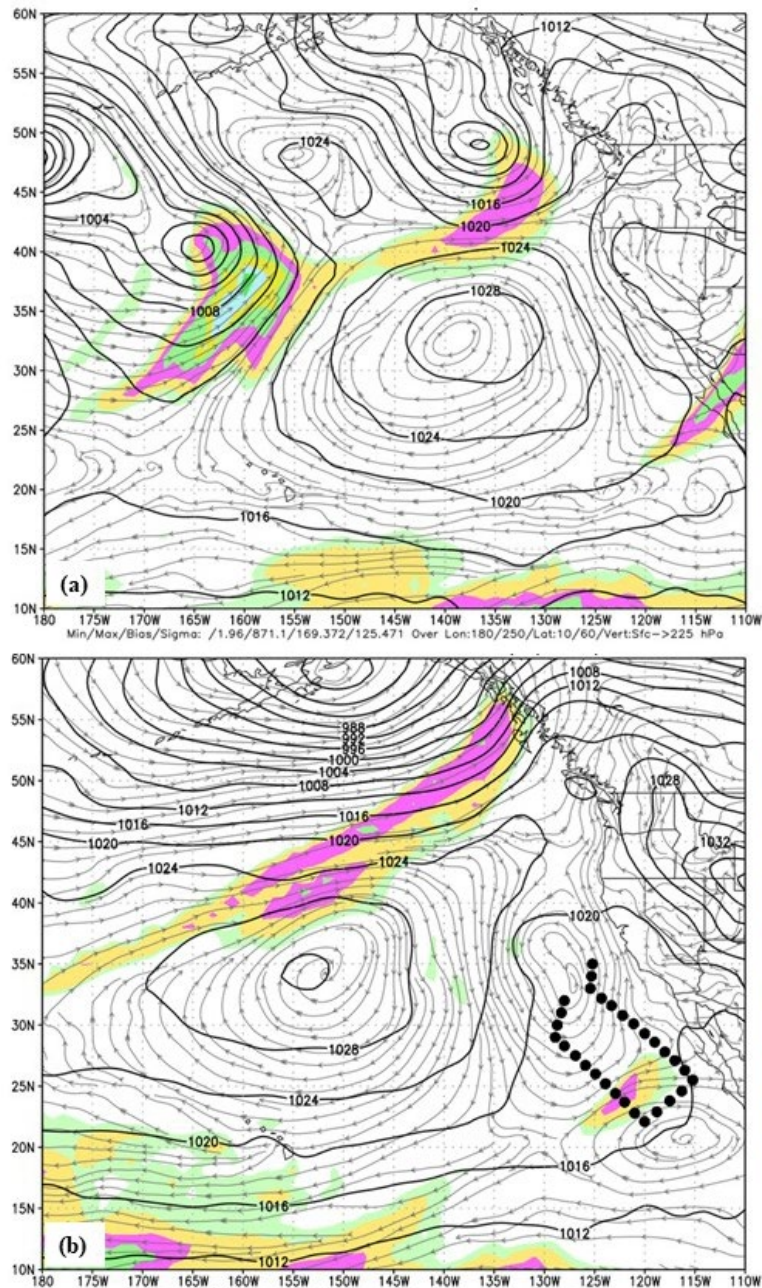


Fig. 16. As in Fig. 3, except for 00 UTC 23 February 2020 (a) and 21 February (b).

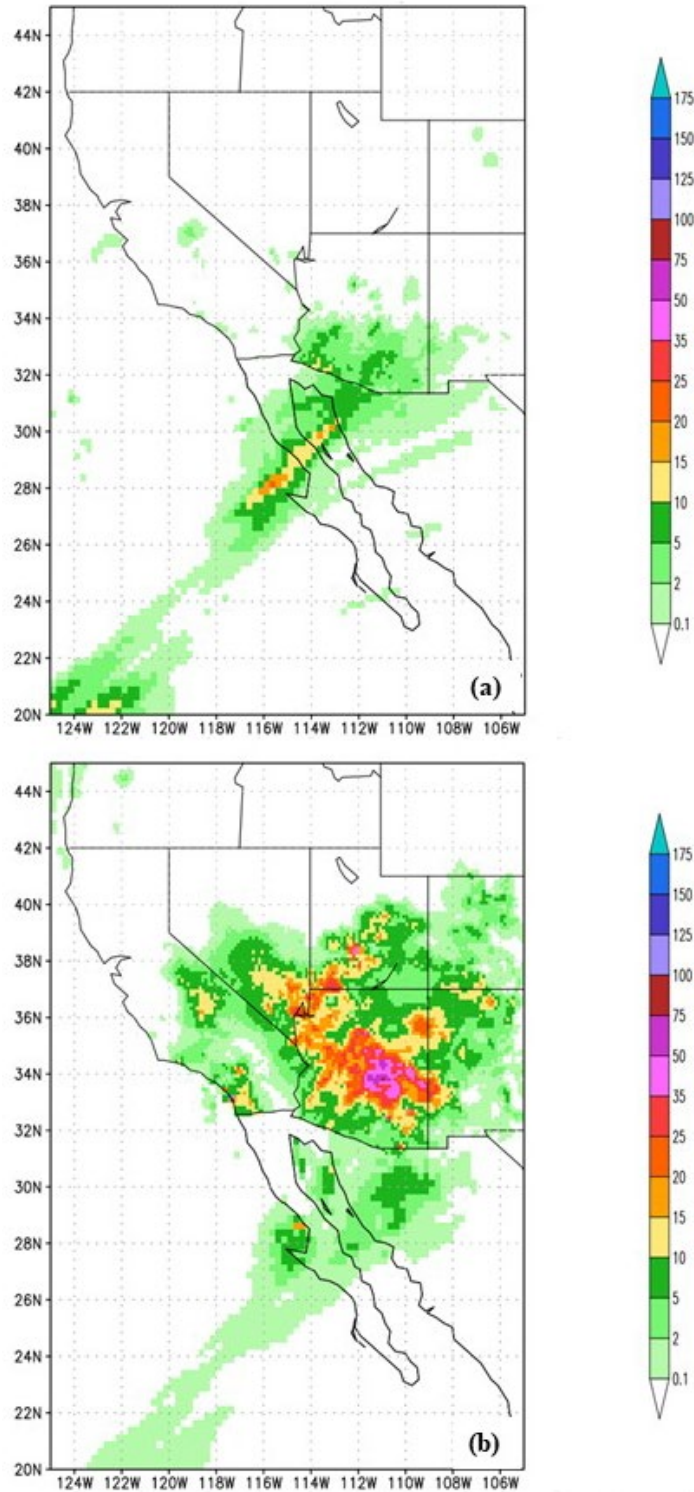


Fig. 17. As in Fig. 5, except for valid dates 12 UTC on 22 February (a) and 23 February (b), and a domain over Baja California, southern California and Arizona.

f. IOP-11: 24 February

The IVT maximum centered at 31-43N, 147-155W (Fig. 18) makes landfall at approximately 06 UTC 25 February with the offshore precipitation maximum at 00 UTC 26 February across southern and southeastern Alaska and northern BC (Fig. 19). There are no land-based precipitation

measurements for this AR event so verification is not feasible. The critical forecast period relative to the observations taken at 00 UTC 24 February is 48 h.

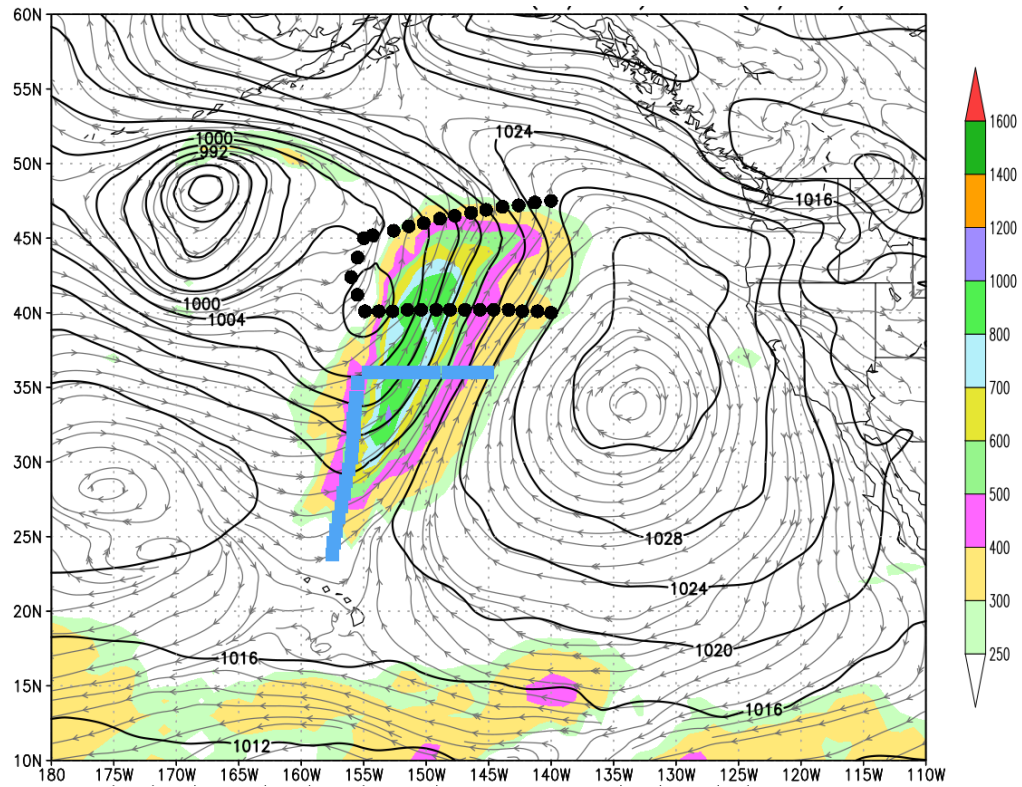


Fig. 18. As in Fig. 3, except for 00 UTC 24 February 2020. Sounding locations from one AFRES aircraft (26, blue) and the NOAA G-IV (30, black) are shown.

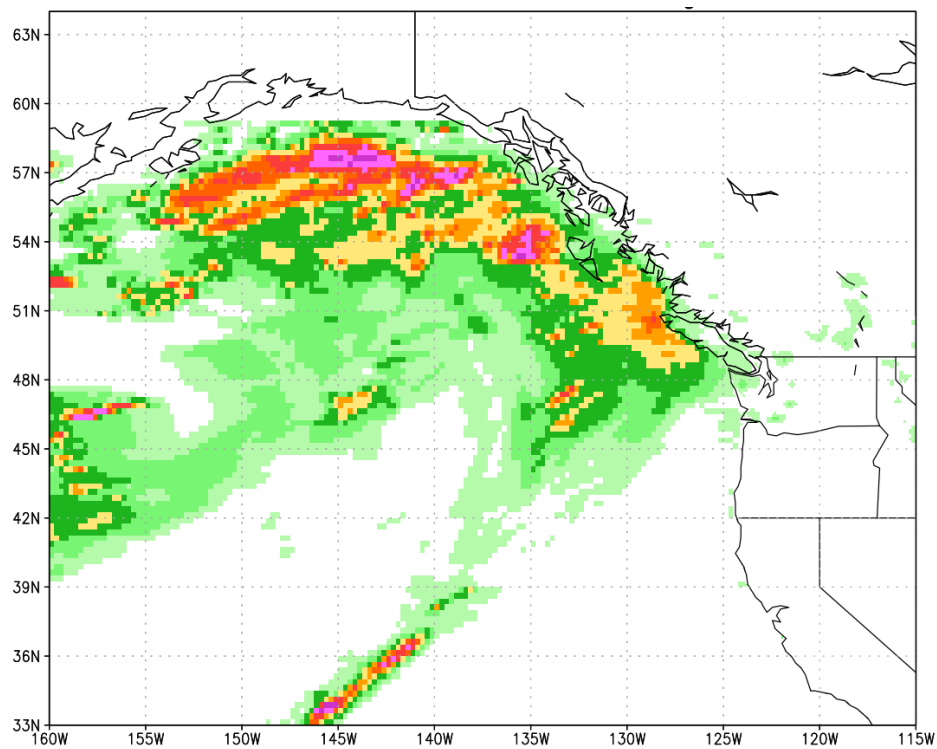


Fig. 19. As in Fig. 11, except for valid date 00 UTC 26 February.

g. IOP-12: 2 March

A very strong AR system developed south of the Aleutians in late February (not shown). By 2 March, the IVT maximum associated with this system has rotated around a high pressure system centered at 40N, 145W and is impinging on the BC coast just north of the USA border (Fig. 20). Major aircraft sorties from both the G-IV and the first AFRES aircraft sampled this IVT maximum, while a second AFRES flight track, centered southwest of a trough over southern California, surveyed another IVT maximum over and the west of Baja California (Fig. 20). Subsequently, over 12 UTC 2 March – 12 UTC 3 March, there is scattered precipitation over southwest California and northwestern Mexico, culminating in heavy rain over southeastern New Mexico and central Texas from 12 UTC 4 March to 00 UTC 5 March (Fig. 21).

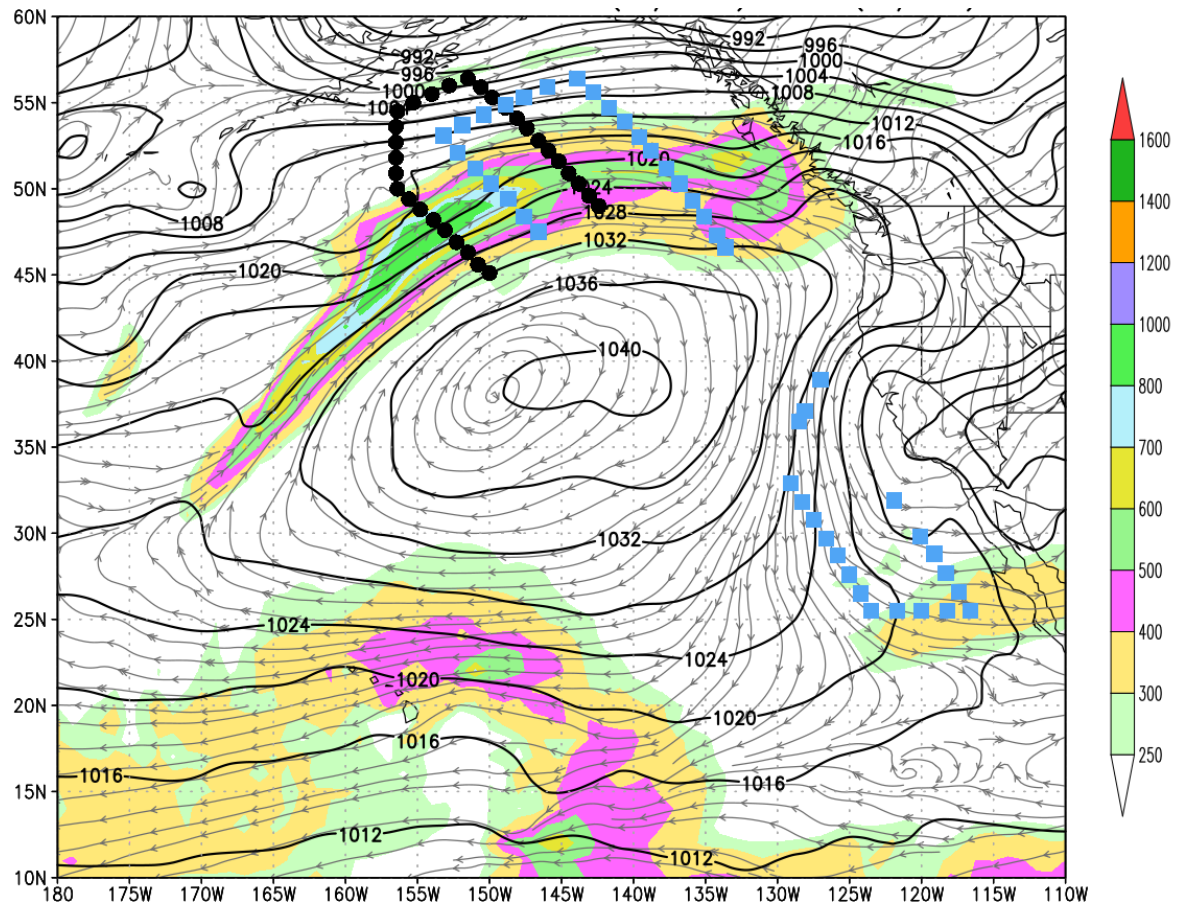


Fig. 20. As in Fig. 3, except for 00 UTC 2 March 2020. Sounding locations from 2 AFRES aircraft (44) and the NOAA G-IV (30) are shown.

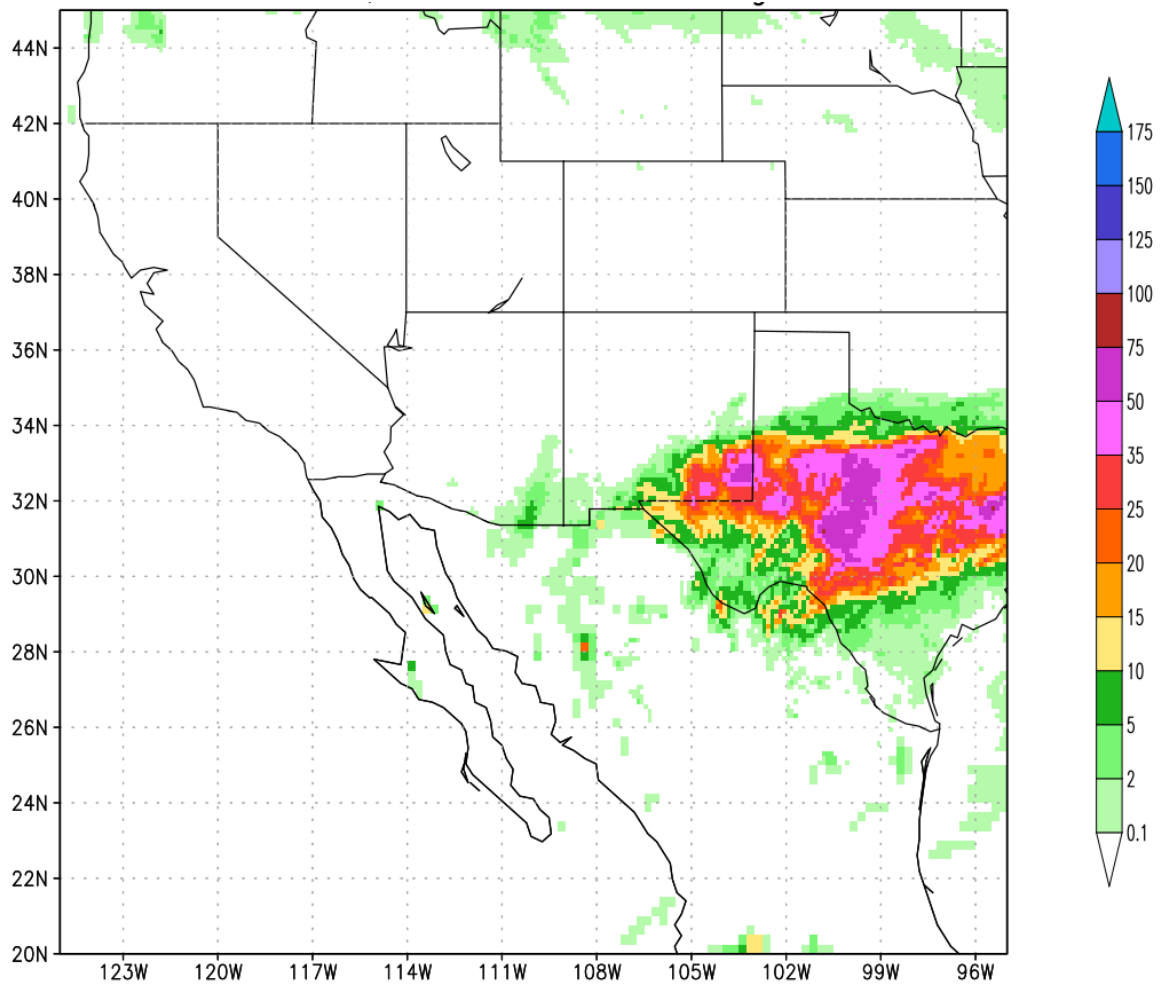


Fig. 21. As in Fig. 5, except for an expanded domain over the southeastern USA and valid date 00 UTC 5 March.

h. IOPs 13-17: 7-11 March

Sorties on these 5 consecutive days were focused on observing two systems with potential impacts over California, Baja California and northwest Mexico. The first system, observed over the 00 UTC 7 March data assimilation cycle (Fig. 22a), was an evolving IVT maxima associated with a developing mid-latitude cyclonic system due west of the Pacific Northwest. Second was an IVT maximum at 20N, 120W, first sampled on 8 March (Fig. 22b) and later sampled at 00 UTC 9-11 March (Figs. 22 c-e). By 12 UTC 9 March, the IVT maxima for these two systems had merged offshore Baja California (not shown). Rain first impacted the San Diego, USA area beginning 12 UTC 10 March and, moving northwestward up the California coast and eastward into both central Arizona and Baja California, it produced significant rainfall through 00 UTC 12 March (Fig. 23a). A second wave of heavy precipitation, again stemming from this stationary IVT maximum, fell over the San Diego area and Arizona from 00 UTC 13 March to 00 UTC 14 March (Fig. 23b-d) and Texas and Oklahoma until 00 UTC 15 March.

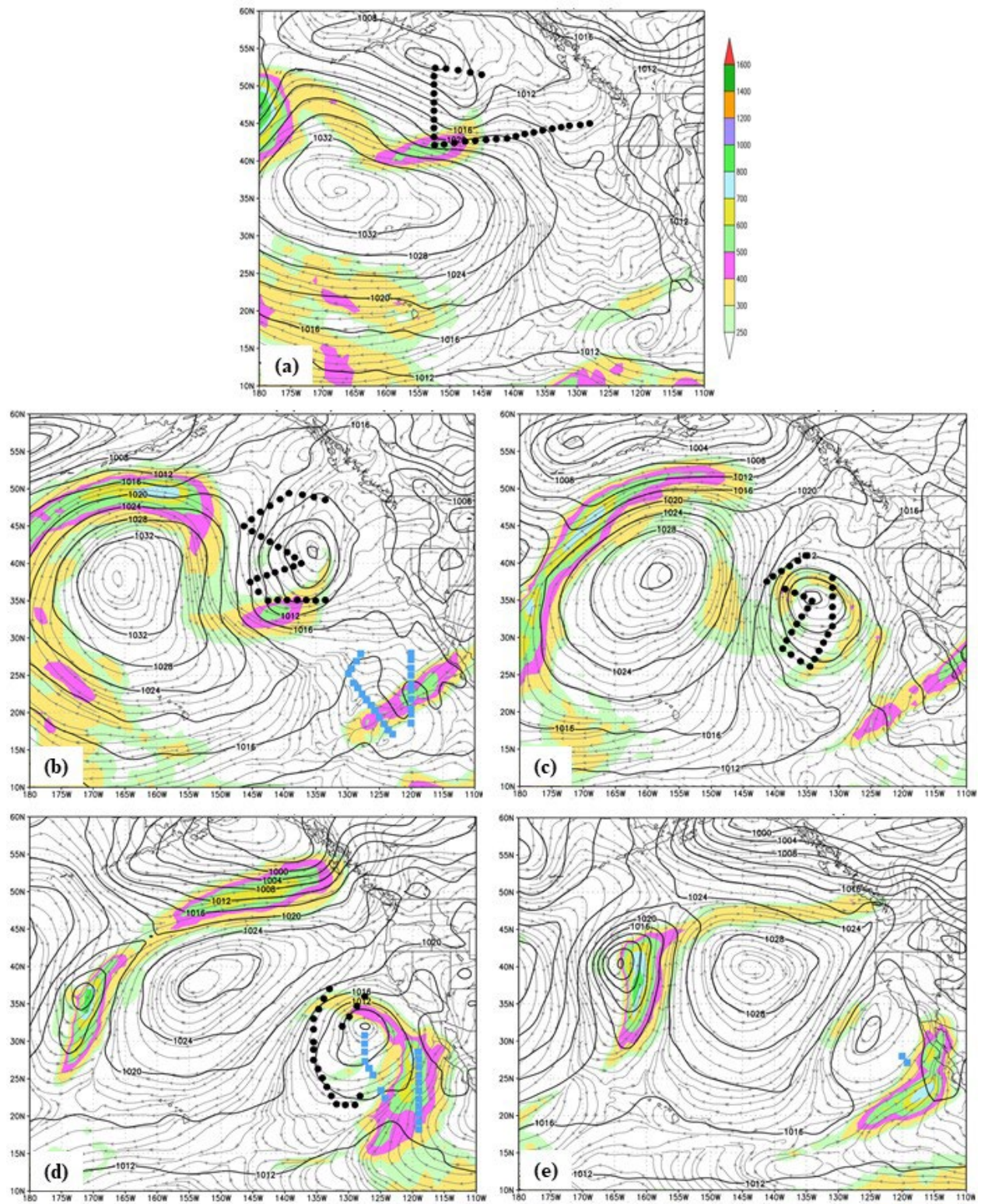


Fig. 22. As in Fig. 3 except for 00 UTC 7-11 March 2020 (a-e, respectively). Sounding locations from the AFRES and NOAA G-IV aircraft assimilated during the 00 UTC cycle are shown for the appropriate dates (Table A1).

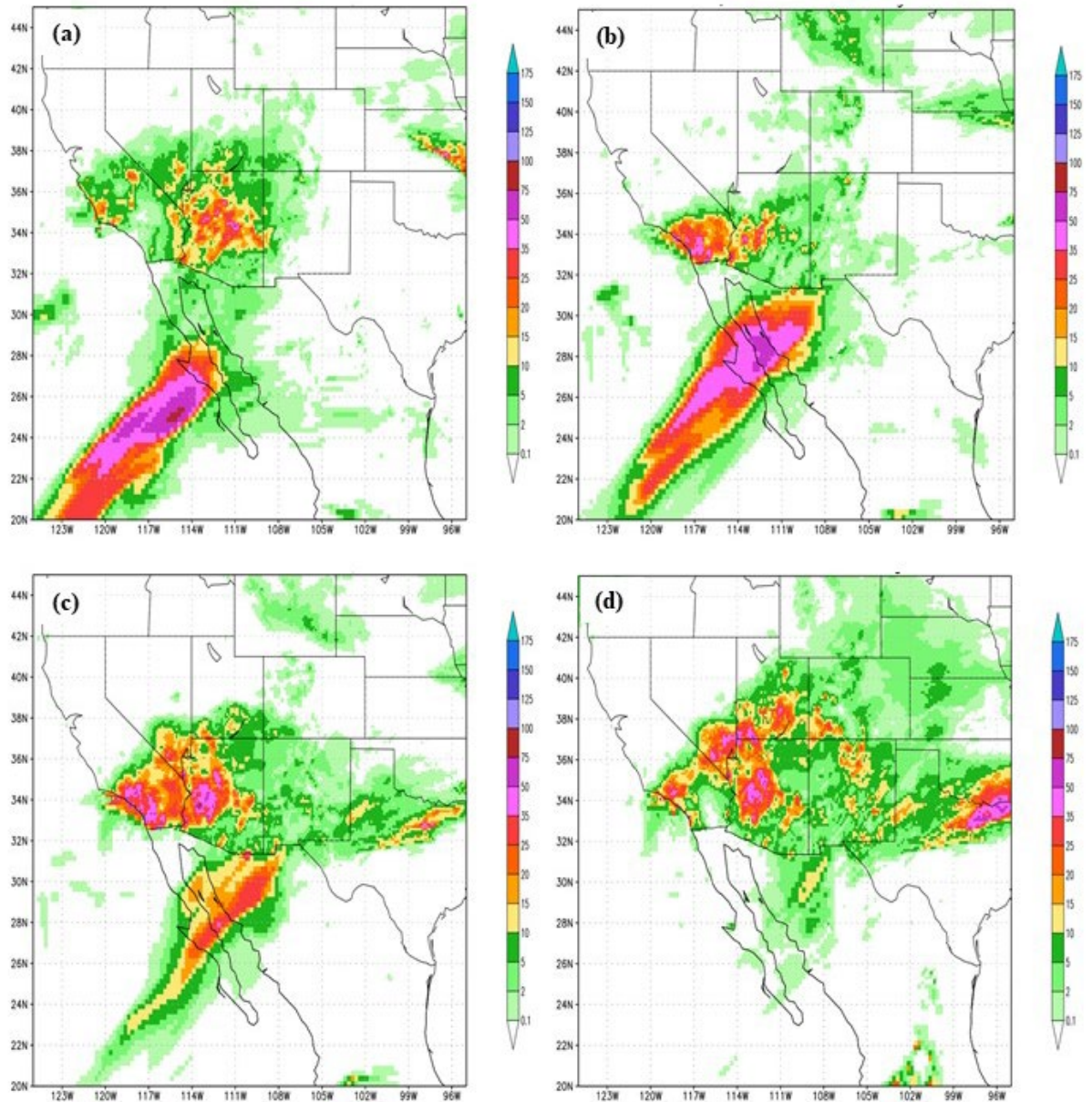


Fig. 23. As in Fig. 21, except for valid dates 00 UTC on 12 March (a), 00 UTC on 13 March (b), 12 UTC on 13 March (c) and 00 UTC on 14 March (d).

5. Results

a. Large-scale mean atmospheric conditions and errors

The time-mean ECMO PMSL analysis over the experimental period and an area encompassing the North Pacific Ocean and most of North America (Fig. 24a) shows a low pressure area centered west of the Aleutians and a Z500 minimum centered over Kamchatka (Fig. 25a). The mid-level trough extends eastward over the northeast Pacific. In the Gulf of Alaska, a secondary PMSL low and co-located Z500 trough are just south of the Alaskan coast. Further downstream are a weak ridge over BC, and a trough over Hudson's Bay extending south through the U.S. Great Lakes.

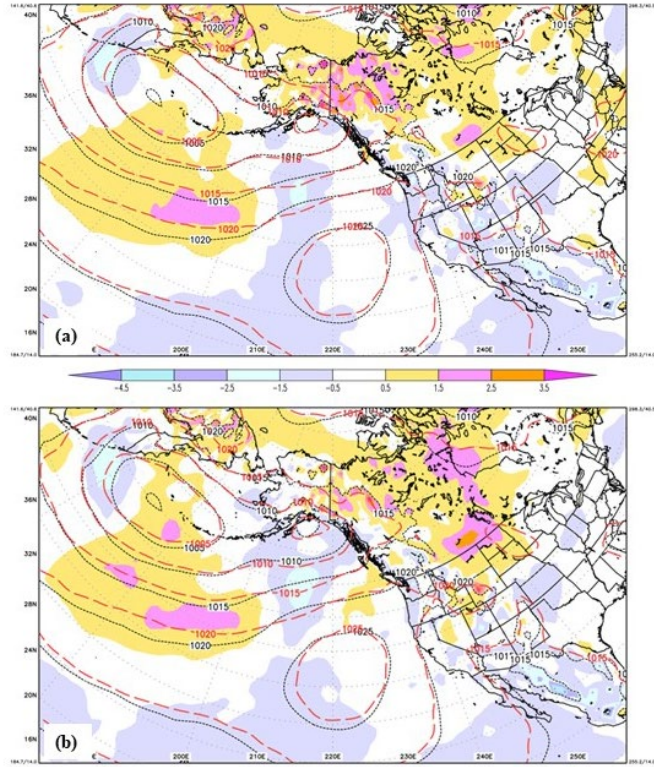


Fig. 24. Time-mean mean sea-level pressure (hPa) over the period 00 UTC 24 January 2020 to 00 UTC 18 March 2020 for the ECMO verifying analysis (black contours) and the CTRL (a) and DENY (b) experiments (red contours). The mean 120 h forecast error (CTRL-ECMO, m) is in solid colors.

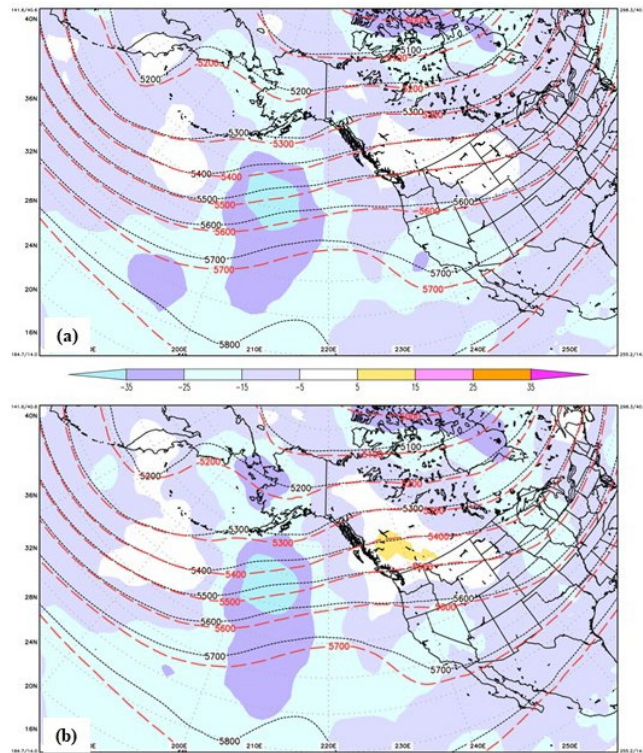


Fig. 25. Time-mean 500 hPa geopotential height fields over the period 00 UTC 24 January 2020 to 00 UTC 18 March 2020 for the ECMO verifying analysis (black contours) and the CTRL (a) and DENY (b) experiments (red contours). The mean 120 h forecast error (CTRL-ECMO) is in solid colors.

An important feature of the mean analyses is the low-pressure over Kamchatka and the Bering Sea. Climatologically, in Northern Hemisphere winter, this area is a major growth area for baroclinic storms. The westward tilt with height of the Bering Sea low pressure center is consistent with the area's frequent cyclogenesis. Cyclones born over the AR2020 experimental period are critical to the formation of ARs impinging on the USA and Canadian west coasts. This winter is, therefore, a highly representative period for investigating ARs and for examining the impact of additional observations to reduce analysis forecast errors in downstream AR landfalls over the USA.

The 120 h CTRL PMSL mean forecast error with respect to the ECMO verifying analysis (Fig. 24a) shows a maximum southeast of the surface low; for Z500 (Fig. 25a) the prominent minimum error is south of the Aleutians at the eastern extent of the mean trough. The DENY experiment (Figs. 24b, 25b) shows very similar patterns over the oceanic domain, with the largest errors shifted slightly poleward compared to CTRL. DENY also has a somewhat larger Z500 error over southwestern Canada. The GFS, therefore, exhibits a notable tendency to propagate troughs too fast to the east with respect to the verification; AR2020 observations do not mitigate this error significantly.

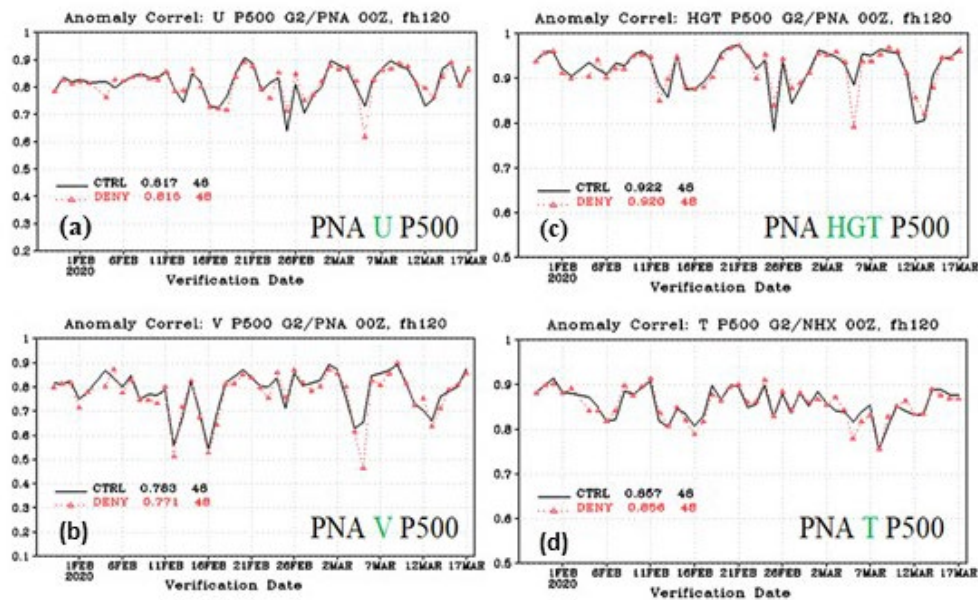


Fig. 26. Anomaly correlation time-series over the experimental period and area 20-75 N and 180-320 E for 120 h CTRL (black) and DENY (red) forecasts at 500 hPa; (a) zonal wind, (b) meridional wind, (c) 500 hPa geopotential height and (d) temperature.

The majority of large-scale, time-mean statistics confirm the lack of significant impact over the North American continent and surrounding oceanic areas (Fig. 26). Mean anomaly correlations for model winds, geopotential height and temperature improved very slightly in CTRL, except for meridional wind, which is 1.2% improved in CTRL. The largest 120 h CTRL improvement is valid at 00 UTC 5 March and initialized at 00 UTC 29 February (Fig. 26), which was 5 days after the nearest IOP.

b. The 00 UTC 29 February case

For the reasons noted above, the 29 February case is clearly an indirect improvement attributable to earlier assimilation of dropsonde data and the non-linear propagation of observational information through the data assimilation system to future initial atmospheric states

as described previously. The CTRL forecast error growth (Fig. 27) is close to that of ECMO out to 72 h, but increases at a higher rate than ECMO beyond 72 h. CTRL error patterns closely resemble DENY out to 96 h, after which the DENY errors increase much more rapidly and error patterns deviate substantially from those of both CTRL and ECMO over the central and northeastern USA and eastern Canada (Appendix C). At 120 h and beyond, DENY forecasts a ridge development over northeastern USA and Nova Scotia instead of a closed low pressure system as indicated in the verifying analysis, resulting in a large anomaly correlation decrease for DENY. More details of the error evolution in this case are presented in Appendix C.

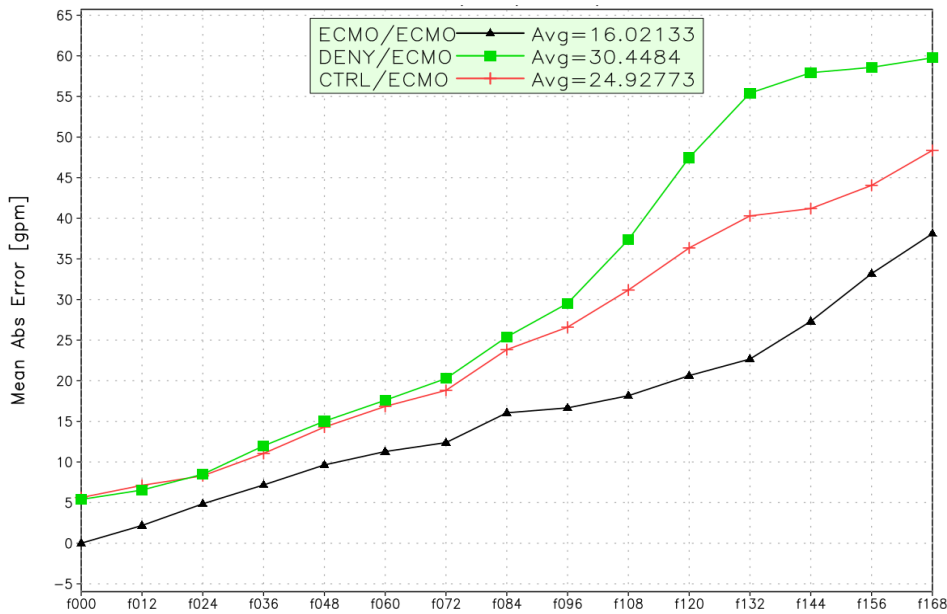


Fig. 27. Z500 MAE for the analysis and 12-168 h CTRL, DENY and ECMO forecasts initialized at 00 UTC 29 February 2020. All MAE are calculated with respect to the verifying ECMO analysis. The verification domain is the same as Fig. C1 (Table 1). The CTRL and DENY analysis MAE results from a Z500 bias between the GFS and ECMO geopotential height fields. The average MAE over the analysis and all forecasts is shown in the legend.

c. Forecast impacts

The direct positive impact on forecast variables is, of course, the purpose of the AR2020 OC. Therefore, to focus on these potential impacts, statistics have been generated for each of the 17 forecasts initialized by AR2020 OC IOP soundings over an appropriate domain (Table 1). For each forecast hour, and for both CTRL and DENY experiments, the statistics are averaged over all 17 IOP cases. Differences between averages are calculated as a percent impact:

$$I(v)=100*(|S_{DENY}|-|S_{CTRL}|)/|S_{DENY}|, \quad (2)$$

where v is a variable (IVT, PMSL, Z, WSPD and SPCH), S_X is a statistic (bias, MAE, min(v), max(v), etc), and $X=CTRL/DENY$ are the experimental runs. $I>0$ reflects an improvement of CTRL over DENY while $I<0$ is a degradation due to assimilating dropsonde soundings. Averaging I over all forecast hours provides an estimate of the total impact for each IOP.

1) IVT impacts

Since the IVT is a strong meteorological feature amid a lower amplitude environment, domain-wide statistics over the area encompassing one or more IVT maxima can evaluate both analysis differences and forecast quality. IVTs propagate through the verification area during the entire

OC so that both direct and indirect impacts are measured by evaluating the IVT errors in CTRL and DENY analyses and forecasts over all IOPs. Since the IVT is an isolated feature, verifications should include only cases in which IVTs are the prominent feature in the verification area, and hence are performed only for IOP dates.

IVT verification statistics (Eq. 2) were generated for 12-168 h forecasts from initial conditions on each IOP date (Table A1). Impacts are measured by the percent improvement (I) of the CTRL MAE relative to the DENY MAE over domain 4 (Table 1). For analyses, I measures the difference between the CTRL/DENY analysis and the ECMO analysis. To measure overall impact, I is averaged over all IOPs for the analysis and each forecast hour. To measure the impact of each IOP, I is averaged over the analysis and all forecasts initialized from that IOP.

(i) *IVT analyses*

DENY and CTRL IVT analyses are compared to ECMO IVTs for all 17 IOPs. Both CTRL and ECMO analyses assimilated OC dropsonde data while DENY analyses did not. For the first IOP, 00 UTC 24 January (Fig. 28a), the DENY analysis has the general shape and magnitude of the ECMO IVT; the DENY-ECMO (D-E) differences are scattered and generally <20% of the IVT magnitude on the leading edge (35 N 145 W) and west-central (W-C, 33 N 155 W) sectors. The CTRL IVT (Fig. 28b) shows the same pattern of CTRL-ECMO (C-E) differences with the ECMO but the W-C sector has larger differences. Isolated maxima (“bullseyes”) are present in both D-E and C-E differences and result from a more variable GFS boundary layer than for the ECMO. IVT analysis differences at some other times also contain bullseyes (not shown).

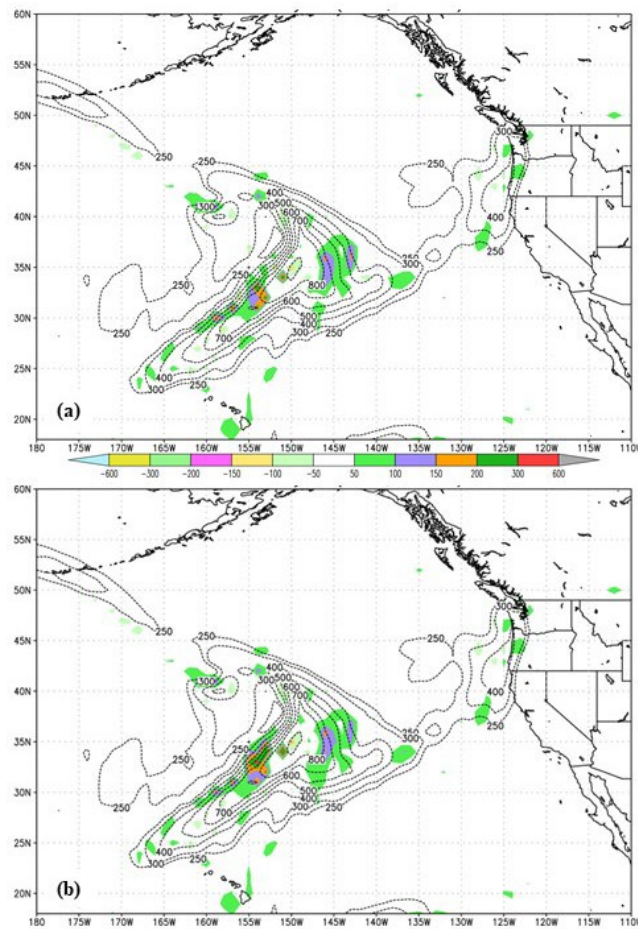


Fig. 28. DENY-ECMO (a) and CTRL-ECMO (b) IVT analysis differences (shaded) and ECMO IVT contours (black) for 00 UTC 24 January 2020.

Overall, there are small, but positive, impacts from the assimilated dropsonde observations on the CTRL IVT structures. The domain-wide mean analysis difference (Fig. 29a) is reduced by 43.8 percent (from -0.740 to -0.416 kg/m/s, i.e., closer to the ECMO IVT mean structure). Maximum C-E values are reduced compared to D-E in 10 IOPs and the IOP average maximum value is decreased by 7.6% (Fig. 29b), while the domain-wide mean absolute difference is reduced by 0.5 percent (not shown). Last, domain-wide minimum C-E values (compared to D-E) are reduced in a minority of the IOPs (7 of the 17) but the IOP-averaged C-E is 6% smaller (Fig. 29c).

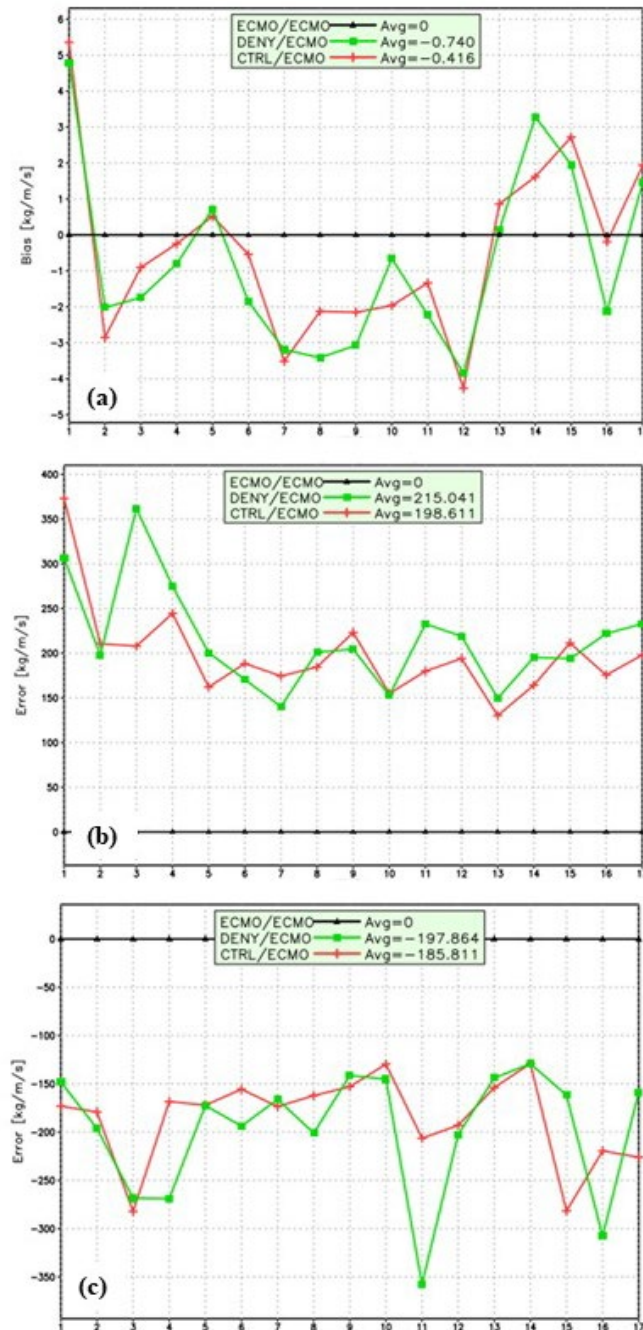


Fig. 29. IVT area-mean (a), maximum (b) and minimum (c) analysis differences [kg/m/s] between DENY (green) and CTRL (red) and the verifying ECMO analysis (D-E and C-E respectively) over the verification domain 4 (Table 1, Fig. 1) for each IOP. The index along the abscissa corresponds to the IOP number (Table A1). Average D-E and C-E values over the 17 IOPs are given in the legend.

(ii) IVT forecasts

The IOP-averaged IVT MAE reduction (Fig. 30) is positive for 24 h forecasts (0.74%), while impacts for analyses, 12 and 36 h forecasts are neutral. Longer forecasts (48-168 h) are mostly degraded, with the largest values at 108 and 168 h. The major contributor to degraded forecasts at 108 h is IOP-13 (00 UTC 7 March, Fig. 31). For each IOP, the average percent error reductions over all analysis and forecast hours are positive for 8 IOPs and negative for 9 IOPs (Fig. 32). The average impact over all IOPs and all forecast hours is -0.47%. The largest improvement is 6.8% for IOP-2 (00 UTC 29 January) and the largest degradation is -11.0% for IOP-13 (00 UTC 7 March). For IOP-2 (Fig. 33a), there are consistent positive impacts both early in the forecast (12-48 h) and later (84-168 h), mostly exceeding 6%. The early impacts are for the observed IVT making landfall while the later impacts are for a new IVT entering the verification domain from the west. For IOP-13 (Fig. 33b), the average reduction is dominated by degradations exceeding 13% from forecasts longer than 72 h, even though there are consistent improvements through 36 h. More details on wind, Z and SPCH errors that are forcing IVT errors are given in sub-section 2 below and in Appendix D.

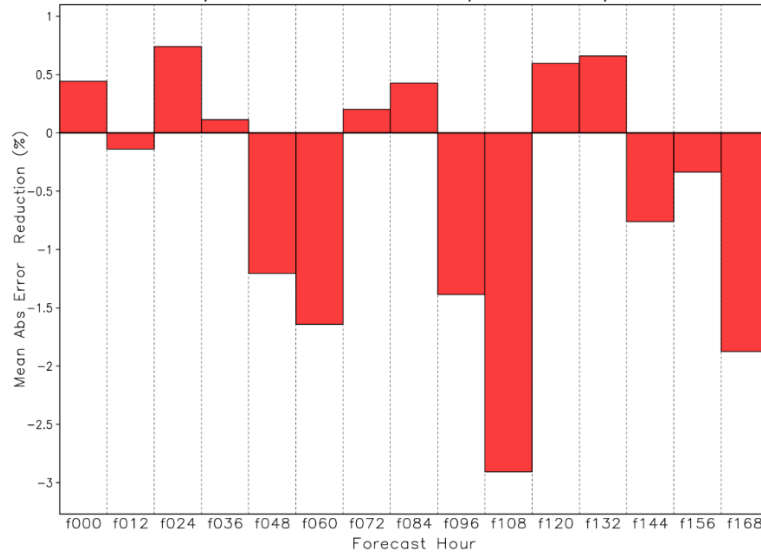


Fig. 30. Average IVT MAE reduction (%) over 17 IOP cases for analyses and 12-168 h forecasts over the domain 18-60N, 180-250 E.

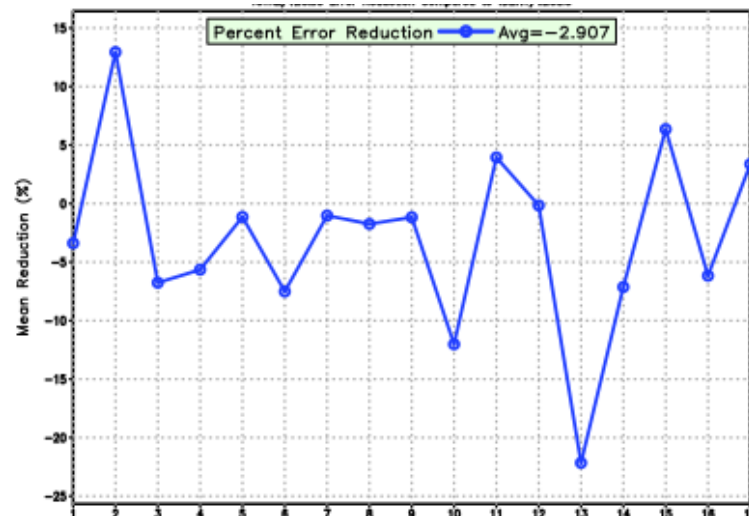


Fig. 31. IVT MAE reduction for each IOP and 108 h forecasts. Abscissa labels as in Fig. 29a.

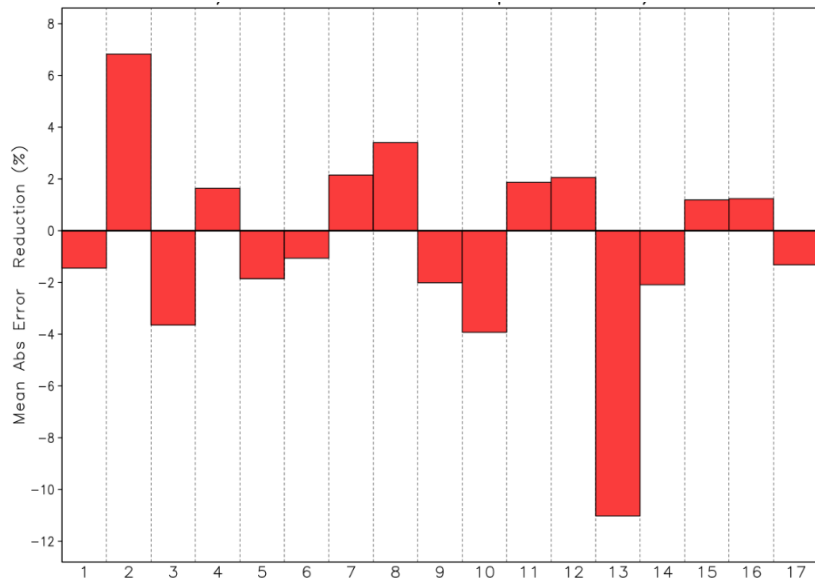


Fig. 32. Average IVT MAE reduction (%) over 0-168 h forecasts for each of the 17 IOP cases.

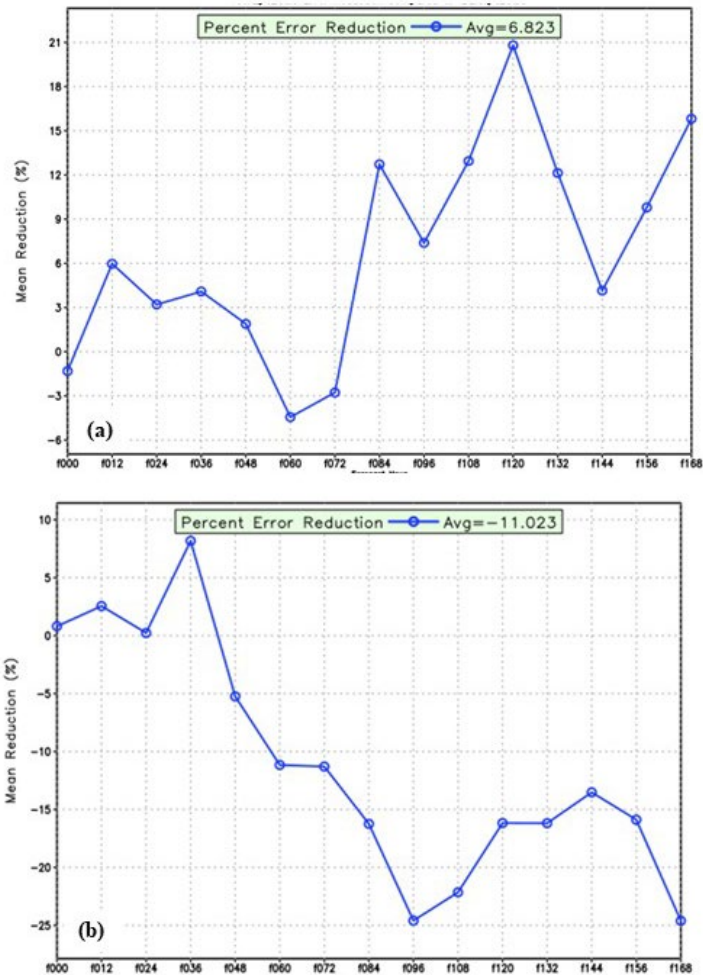


Fig. 33. IVT MAE reduction for 00-168 h for IOP-2 (a), initialized at 00 UTC 29 January and IOP-13 (b), initialized 00 UTC 7 March.

(iii) IVT impact summary

To summarize all analysis and forecast impacts, I is categorized by assigning an integer to a range of impact magnitude (Fig. 34a). Over the complete set of IOPs and forecast ranges, the categorized I (Fig. 34b) varies over the full range ($-4 \leq I \leq 4$), generally without strong coherence from one IOP to the next or from one forecast projection to smaller or larger projections. For 36–60 h forecasts, there is a weak trend from small negative impact to moderate positive impact for consecutive IOPs (14–16 February and from negative impact to moderate positive impact (7–10 March). However, for consecutive IOPs on 4–6 February, almost all impacts are negative for 12–108 h.

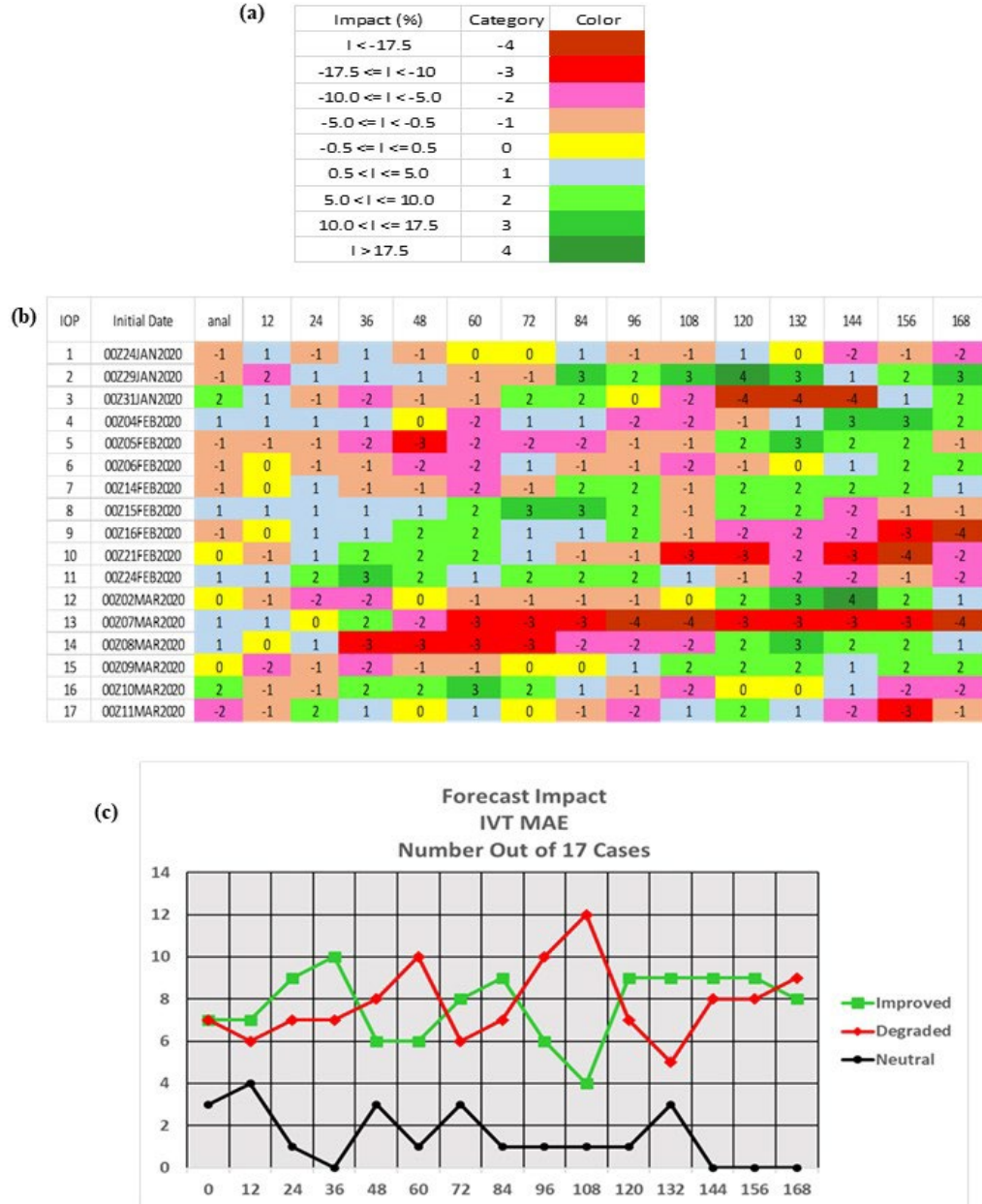


Fig. 34. IVT MAE impact (I , %) categories for CTRL error improvement relative to DENVY (a), categorical summary (b) for each IOP and forecast hour (analysis and 12–168 h), and number of IOPs (c), out of 17, improved (green), degraded (red) or with neutral impact (black). Positive categories indicate improvement, negative are degradations and zero (magnitude of 0.5% or less) is neutral.

From the analysis to 108 h forecasts, there is an expected downward trend in the number of improved IOP cases and a corresponding positive trend in the number of degraded cases (Fig. 34c). Also expected is the number of neutral cases dwindling to zero beyond 132 h. The number of improved IVT cases beyond 108 h is unexpected, as is the corresponding improvement to both wind speed and moisture forecasts over the same forecast period as shown later.

For a particular IOP, as noted for IOP-2 above, a common, major meteorological feature over the AR OC period is a subsequent IVT system entering the verification domain, typically from the west or southwest. Of course, this system is not observed directly, but its performance can be impacted indirectly (positively or negatively) by previous observations through the GDAS analysis and cycling. Ideally, impacts would be positive but, practically speaking, limited-area observations have the most chance of improving downstream, rather than upstream, forecasts. However, some of the impact at longer forecast projections may be positive overall, such as the IVT MAE at 120-132 h (Fig. 30). This may also be the case for some or all of the model forecast variables themselves as noted below.

2) PMSL, Z, WSPD and SPCH impacts

(i) IOP-averaged impacts

The overall MAE percent improvements for PMSL, averaged over all analyses, forecast hours and IOPs, and for Z, WSPD and SPCH (each additionally averaged over 200-925 hPa) are -0.99, -0.70, 0.23 and 0.56 respectively. The approximately 1% degradations for PMSL and Z are consistent as expected, while the overall improvements for WSPD and SPCH are encouraging. These average percent improvements, however, show varied impact with height (Fig. 35). Degradations in Z dominate the mid- and lower-troposphere, while WSPD is positive (0.1-1.2) at all levels up to 400 hPa and SPCH improvements are largest (~1%) up through the lower- and mid-troposphere before decreasing in the upper-troposphere.

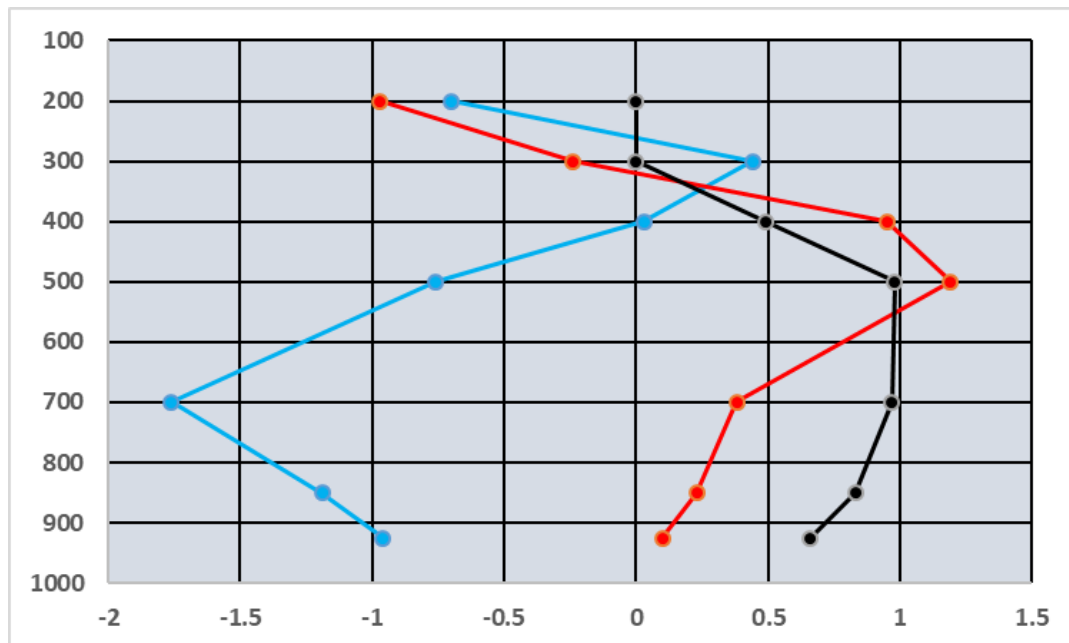


Fig. 35. Vertical profile of MAE improvement (%) averaged over all analyses, forecast hours and IOPs for geopotential height (blue), wind speed (red) and specific humidity (black). Specific humidity improvements are zeroed out at 200-300 hPa due to insufficient precision in their calculation.

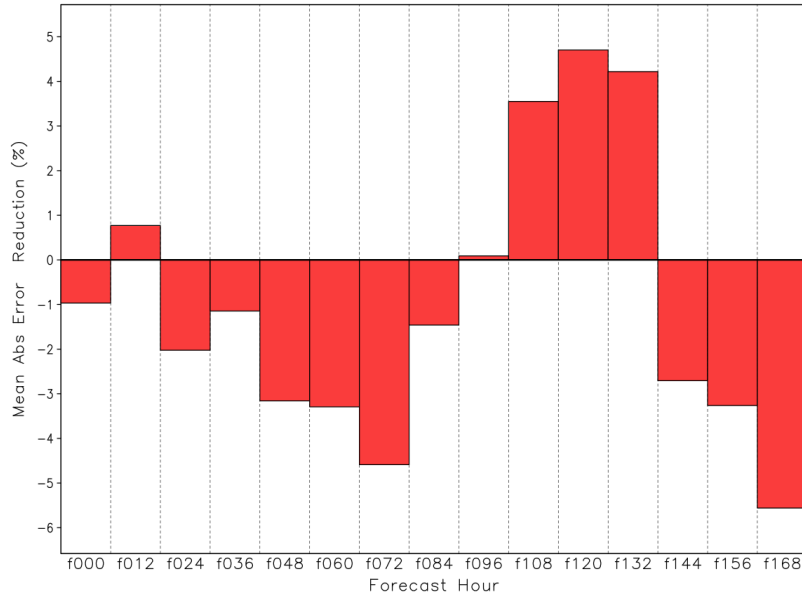


Fig. 36. Average MAE reduction (%) over 17 IOP cases for 12-168 h mean sea-level pressure forecasts, verified against ECMO analyses. Positive reductions indicate improvement for CTRL relative to DENY. The verification domain is Domain 3 (35-65 N, 200-245 E).

The IOP-averaged PMSL MAE for the analysis and 12-168 h forecasts, verified against ECMO fields (Fig. 36), shows less than 1% improvement at 12 h, 2-5% CTRL degradation from 24-72 h, 3-5% improvement at 108-132 h, and 3-6% degradation from 144-168 h. CTRL geopotential height analyses (Fig. 37) are closer to the verification, and a similar pattern to PMSL forecasts is shown for all levels over the entire forecast range: CTRL MAE improvements for 12-36 h, degradations for 48-96 h, 1-4% improvement at 120 h and 1-8% degradation at and beyond 132 h. Impacts at 200 hPa are generally the largest magnitude but lower and mid-tropospheric impact magnitudes are larger at 156-168 h. The magnitudes of IOP-averaged PMSL and Z error improvements over the entire forecast range are not highly correlated with IVT errors (Figs. 36, 37, 30).

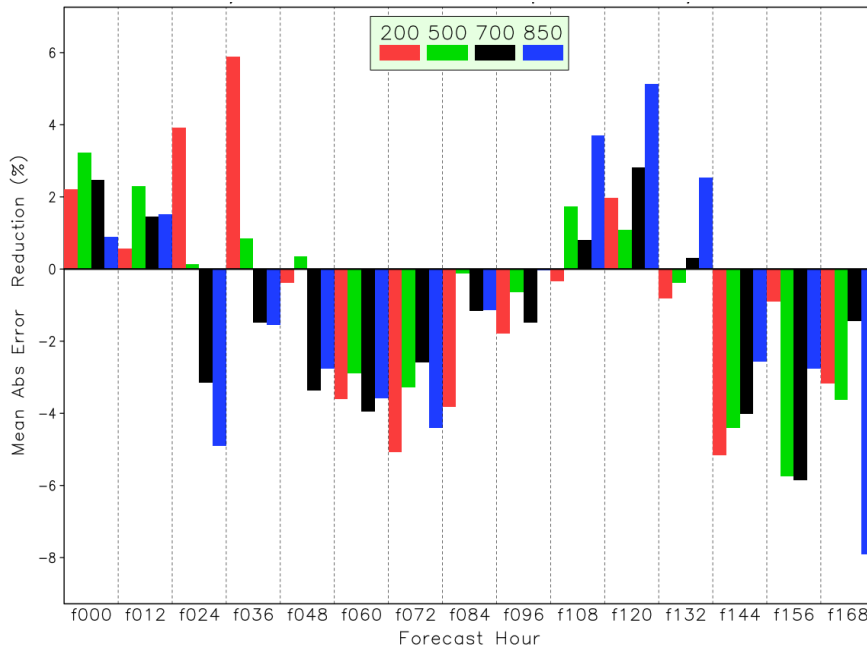


Fig. 37. As in Fig. 36, except for geopotential height at 200, 500, 700 and 850 hPa.

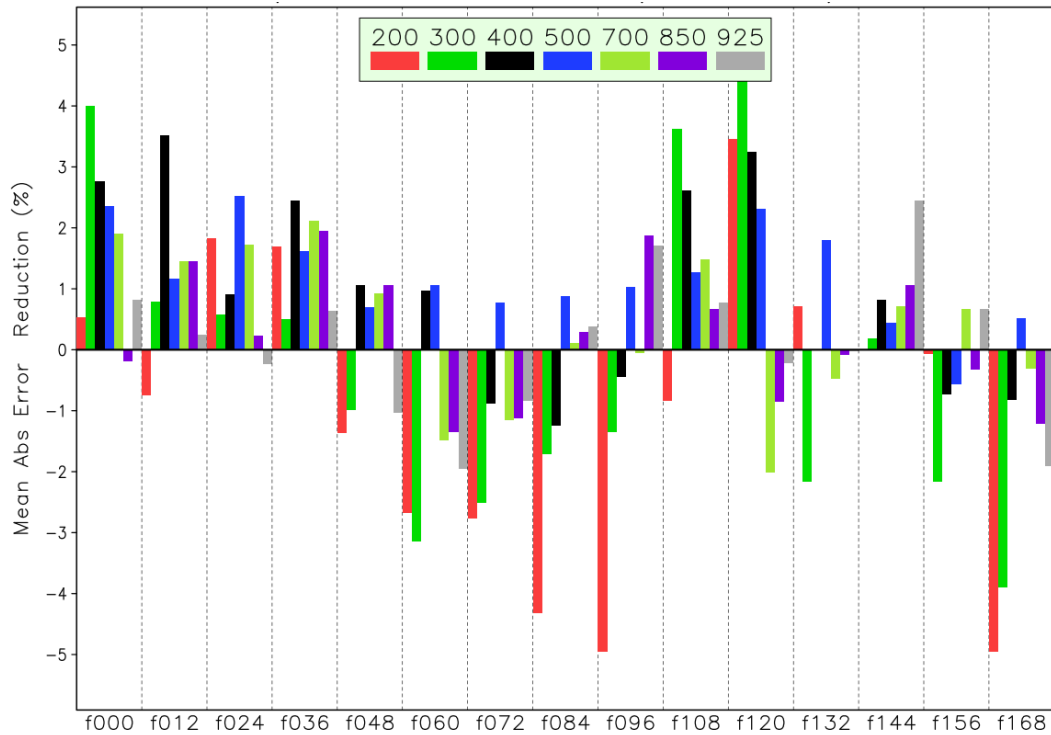


Fig. 38. As in Fig. 36, except for wind speed at 200, 300, 400, 500, 700, 850 and 925 hPa. Missing data, e.g. large values removed due to division by near-zero averages, are represented as a zero value.

Wind speed impacts (Fig. 38) are generally 1-4% improvements of CTRL over DENY across all pressure levels out to 48 h. Beyond 48 h, impacts at 200-300 hPa have the largest magnitudes and dominate the negative impacts from 48-96 h. Average negative wind impacts for almost all levels at 60-96 h appear to precede negative impacts on IVT forecasts at 96-108 h, while positive wind impacts at 108-120 h tend to precede positive IVT forecast impacts at 120-132 h (Fig. 30).

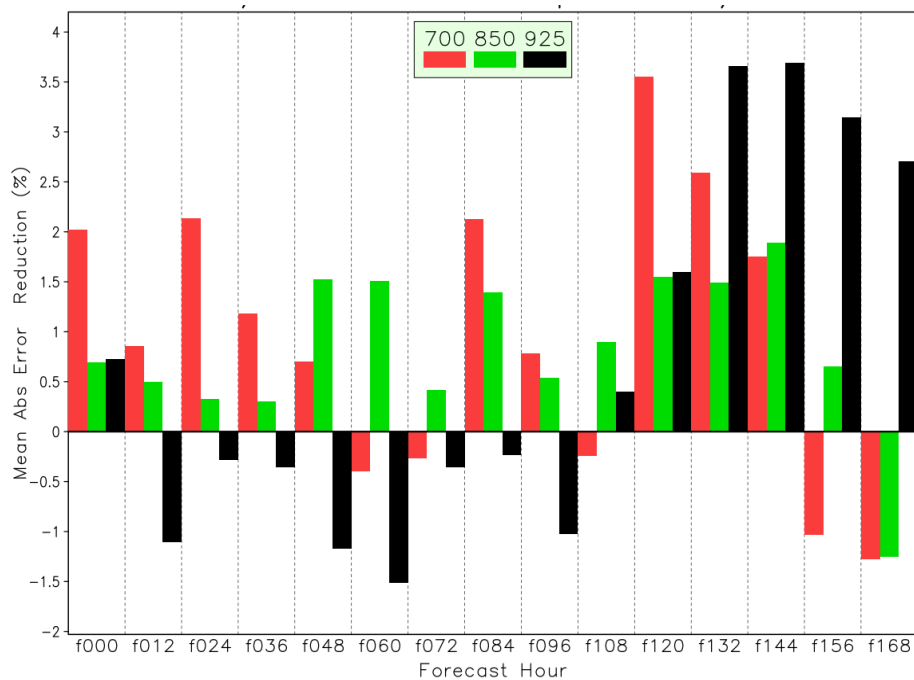


Fig. 39. As in Fig. 36, except for specific humidity at 700, 850 and 925 hPa.

Specific humidity impacts (Fig. 39) for 700, 850 and 925 hPa are also generally positive, except at 925 hPa for 12-96 h forecasts. These impacts appear to be robust: for all analyses over the 700-925 hPa layer, improved moisture MAE for CTRL vs DENY occurs in 65% of the 17 IOPs, and positive forecast impacts occur in 53% of all 12-168 h forecasts in these layers. These positive impacts range from 47% of all IOPs at 12-24 h to 63% at 120 h and 144 h and are more consistent across all forecast ranges than for geopotential height and wind speed. Overall, they confirm that the AR2020 OC dropsonde specific humidity observations are uniquely accurate, high resolution measurements that provide consistent positive value to AR forecasts by improving the initial CTRL moisture analyses and by improving the majority of CTRL moisture forecasts relative to the DENY experiment.

(ii) Forecast-averaged impacts

By averaging the percent impact for the basic model forecast variables over all forecast hours, as done for IVT (Fig.32), the cumulative impact of added observations from each IOP can be assessed. For PMSL (Fig. 40a), the largest positive (+12.9%) and negative (-14.3%) impacts are for IOP-2 (29 January) and IOP-13 (7 March) respectively. PMSL impacts for other IOPs range from +6.9 to -7.9%, but with 9 of the remaining 15 IOPs having less than 3% impact magnitude. Across all IOPs, PMSL impacts do not always appear correlated with those for Z200 and Z500 (e.g., IOP-2, Figs. 40a, b), but there is somewhat higher correlation with Z700 and Z850. The large PMSL degradation for IOP-13, followed by more neutral impacts in subsequent IOPs 14-17, is also reflected in all geopotential height impacts and implies successive forecast improvement in consecutive IOPs through the GDAS cycle. Tropospheric wind speed impacts across all IOPs (Fig. 40c) appear well correlated with PMSL impacts. Lower tropospheric moisture impacts, while generally positive, are less correlated with the other model variables (as expected), but nevertheless show strong correlation with winds and PMSL for IOP-2 and IOP-13.

While the forecast-averaged PMSL impact for IOP-8 (15 February) is close to zero (Fig. 40a), impact over the full forecast range (Fig. 41a) varies from negative (12-48 h), to positive (60-132 h) and finally large negative (144-168 h) as the forecast proceeds. IOP-9 (16 February) also has a similar pattern (Fig. 41b) with small positive impact (12-24 h), followed by moderate negative impact (36-84 h), positive impact (96-132 h) and large negative impacts at 144-168 h. Impacts for these IOPs are discussed further below and, in more detail, in Appendix D.

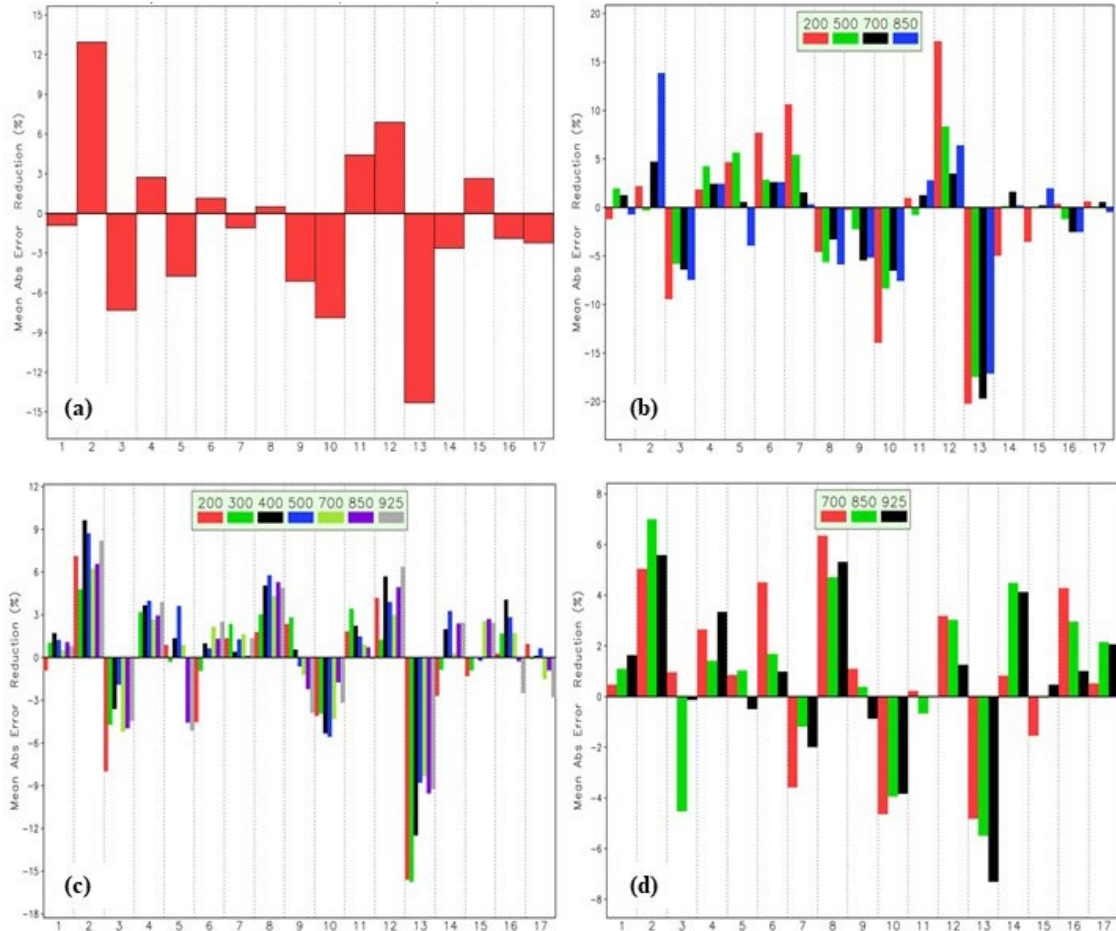


Fig. 40. As in Fig. 11, except for mean sea-level pressure (a), geopotential height at 200, 500, 700 and 850 hPa (b), tropospheric wind speeds (c), and specific humidity at 700, 850 and 925 hPa (d).

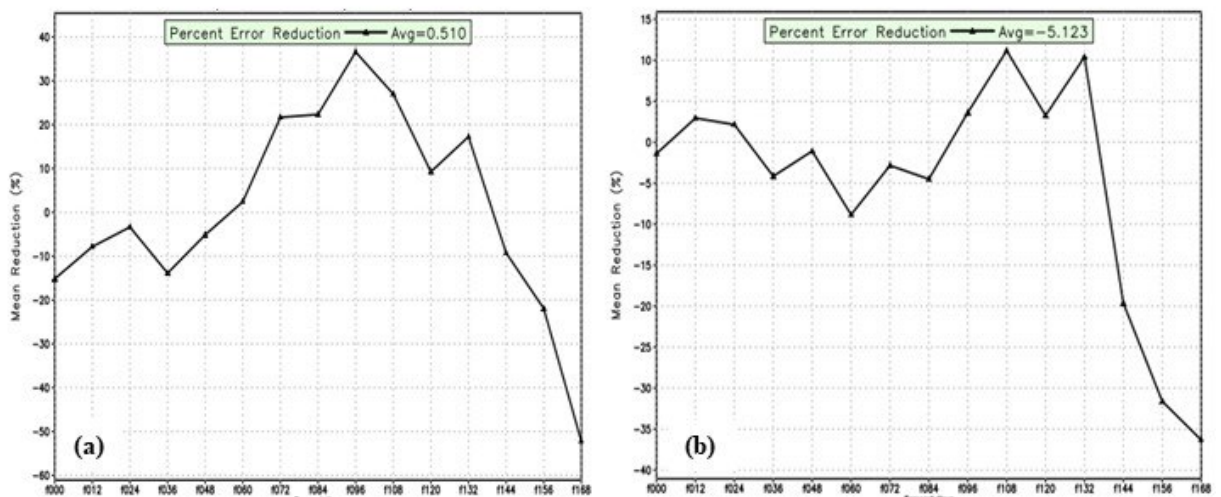


Fig. 41. Mean sea-level pressure MAE reduction for 00-168 h for IOP-8 (a), initialized at 00 UTC 15 February and IOP-9 (b), initialized at 00 UTC 16 February.

(iii) Impacts for IOPs 2, 8-9 and 13

The positive impact for IOP-2 occurs for wind speed and moisture (both critical components of IVT) at the selected levels (Figs. 40c, d), and also for lower tropospheric geopotential height, but to a lesser extent for Z200 and not for Z500 (Fig. 40b). Forecast-averaged geopotential height impacts over the deep 200-925 hPa layer for IOP-2 average 6.1%, with a range of -0.3% at 500 hPa to 15.2% at 925 hPa (not shown) and improvements of 2-4% above 500 hPa. This vertical structure suggests that the AR OC dropsondes and the GDAS provide an improved analyzed CTRL baroclinic structure for this IOP and that both observations and data assimilation play a positive role in forecast improvement. From the analysis time (00 UTC 29 January) to the end of the major rainfall event over Washington State at 00 UTC 2 February (Figs. 6-9), the forecast CTRL wind speed, moisture and IVT verifications are generally improved over the DENY experiment.

While IOP-8 (15 February) is characterized by a small positive (3.4%) impact on IVT (Fig. 32), and overall positive impact for both wind speed (~4%, Fig. 40c) and specific humidity (~4%, Fig. 40d), the impact for mean sea-level pressure is near zero (as noted previously) and negative (~-4%) for geopotential height (Fig. 40b). This neutral/negative impact on the mass field is exemplified by the Z700 errors (Fig. 42), and is driven by a rapid increase of the CTRL IOP-8 forecast error from 108 h, when the CTRL was improved, onward to 168 h, when there is a major CTRL degradation as the CTRL error is double that of the DENY experiment. A similar negative forecast impact at 132-168 h for geopotential height is present throughout the 200-925 layer for IOP-8 and also for IOP-9 (16 February).

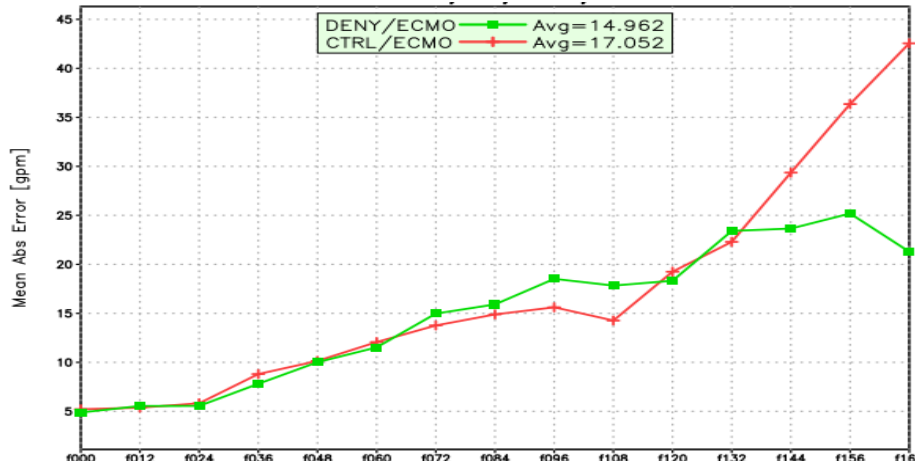


Fig. 42. Z700 forecast MAE for IOP-8 (15 February).

Enhanced IVT degradation at longer forecast times often reflects errors for IVTs that enter the verification domain during the forecast, typically from the western boundary (see Section 4 for examples). However, the large geopotential height CTRL degradations at 144 h onward for IOPs 8-9 derive primarily from enhanced errors in the CTRL large-scale mass field due to increased growth of a low pressure system over the northern Gulf of Alaska and Prince William Sound and progressive southeast propagation of this system to west of Vancouver Island subsequent to IOP-9. While the DENY experiment has errors, they are less impactful on the day 6-7 forecast accuracy so that the CTRL error growth unexpectedly produces an inferior forecast.

The IOP-13 forecast includes two major mid-latitude IVT systems and a long-lasting sequence of IVTs landfalling over Baja California and southern California. The first mid-latitude IVT enters the verification domain at the analysis time (00 UTC 7 March, Fig. 22a) and the second develops on the southwestern end of the previous IVT (Fig. 22c) on 06 UTC 9 March in association with a

developing cyclone at 30 N 178 W. This cyclone continues to strengthen as it moves north northeast (Figs. 22d-e) to 42 N 165 W by 12 UTC 12 March.

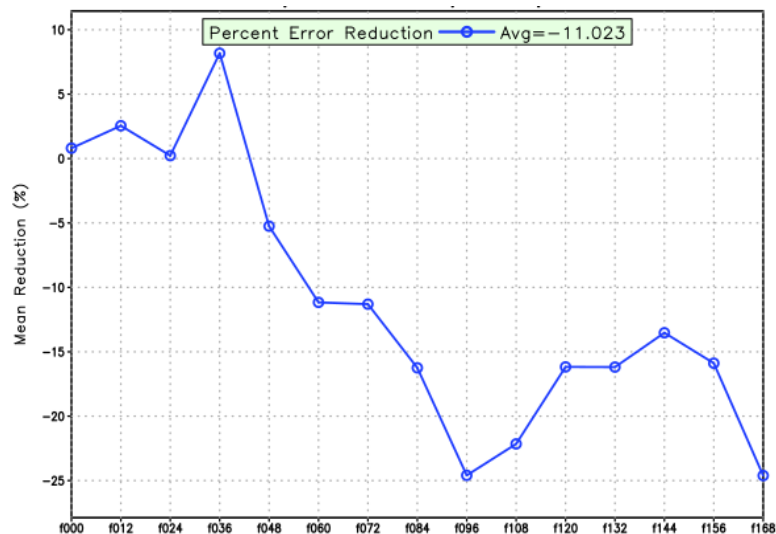


Fig. 43. IVT MAE reduction for 00-168 h for IOP-13 initialized at 00 UTC 7 March.

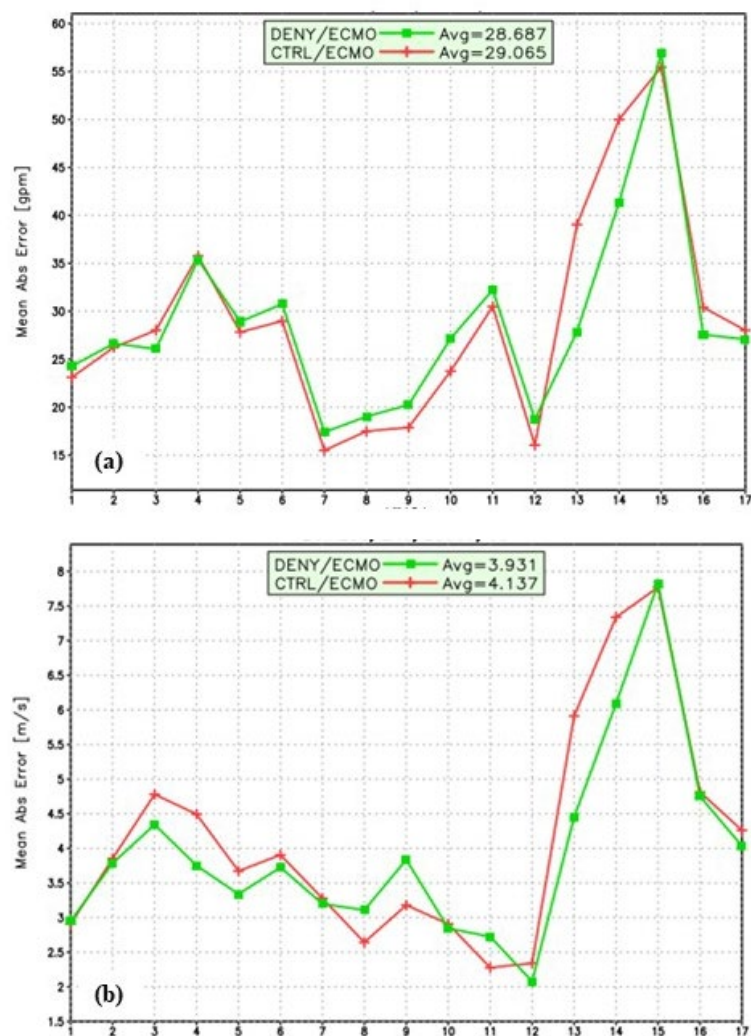


Fig. 44. Z500 (a) and WSPD200 (b) MAE for 84 h forecasts and each of the 17 AR2020 IOPs.

In the first 36 h of the IOP-13 forecast, the CTRL has neutral to positive IVT MAE reductions of 2% at 12 h and 8% at 36 h. However, the forecast rapidly degrades thenceforth to a major degradation of 25% by 96 h (Fig. 43). Similar impacts occur for PMSL, geopotential height and wind speed (not shown), resulting in large negative forecast-averaged impacts for all variables (Figs. 40a-d). For example, at 84 h, IOP-13 and IOP-14 CTRL and DENY forecast errors for Z500 and WSPD200 are almost twice as large as for other IOPs (Figs. 44a, b) and they, along with corresponding errors at other levels, account for the degradation in the IVT landfall errors at 12 UTC 10 March (Figs. D9a-c)

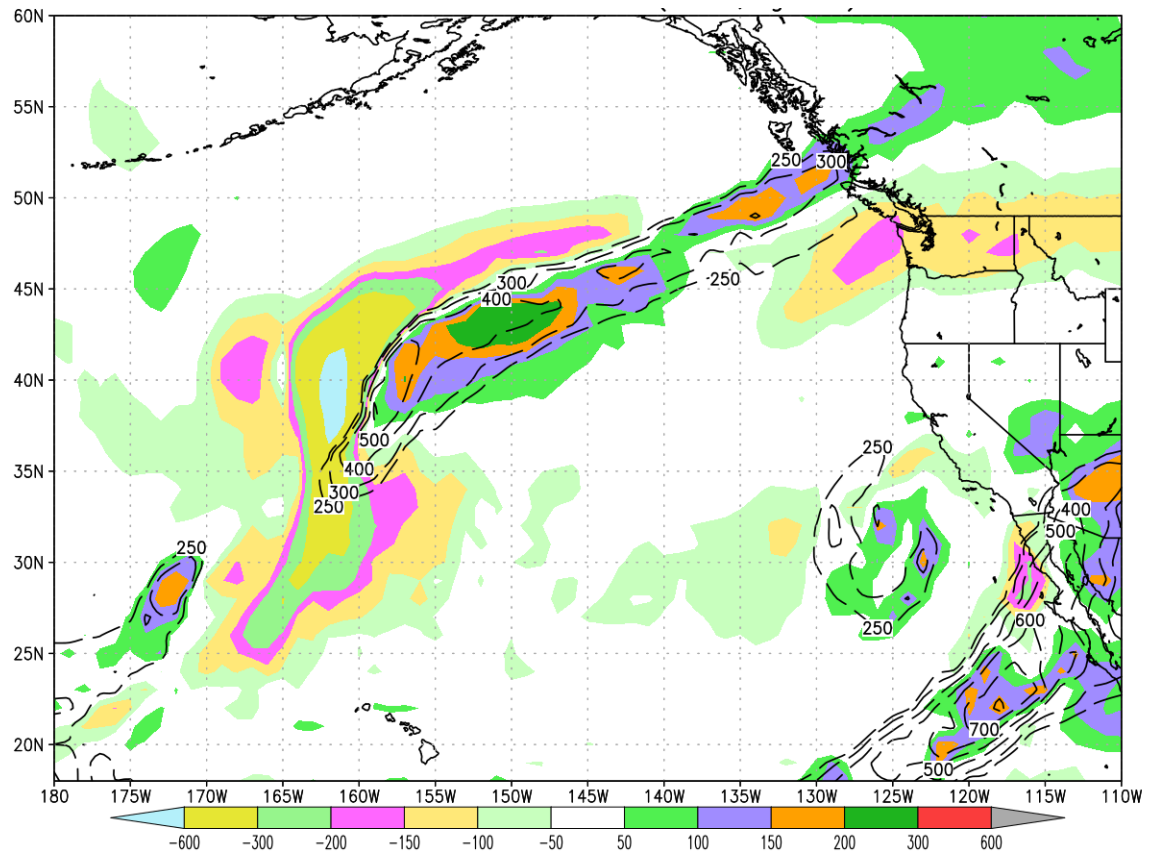


Fig. 45. IOP-13 (7 March) CTRL IVT 96 h forecast error (shaded) and forecast contours (black) valid for 00 UTC 11 March 2020.

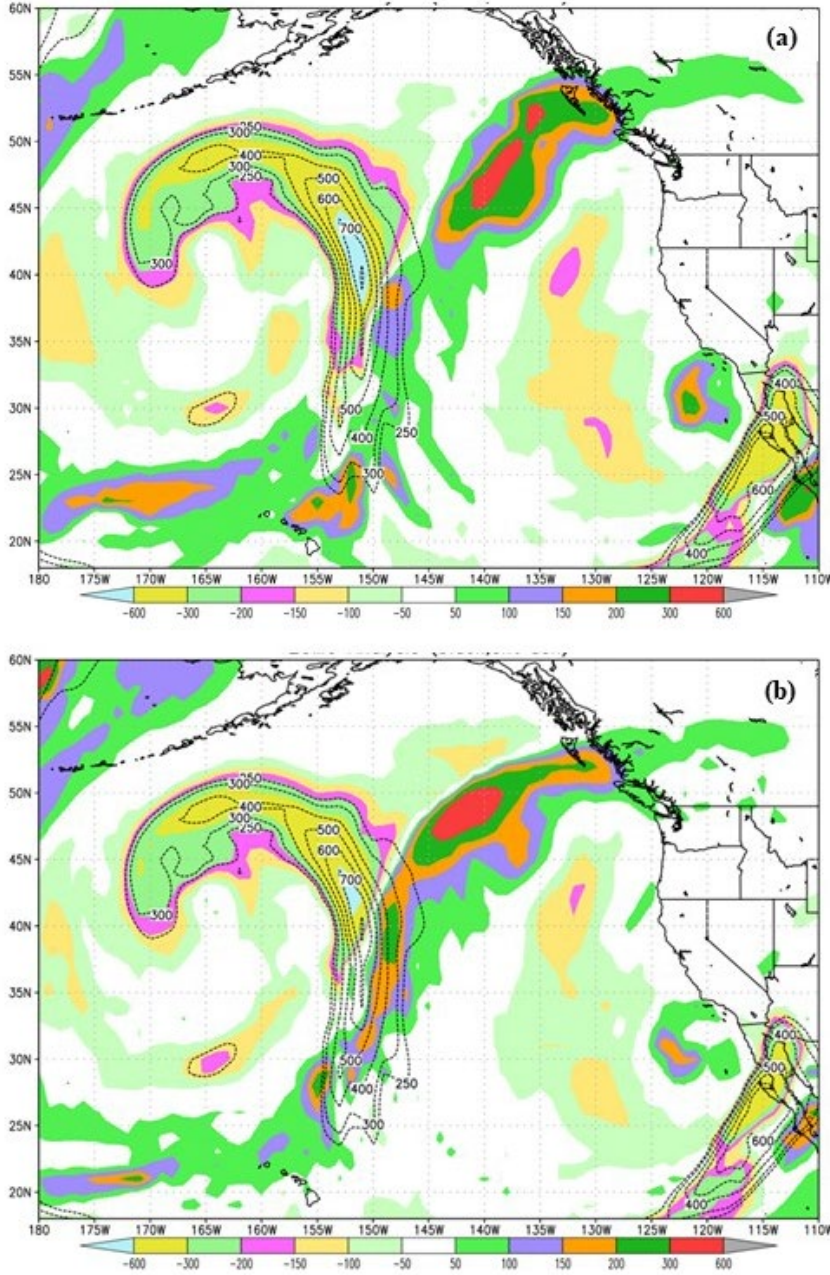


Fig. 46. IOP-13 (7 March) CTRL (a) and DENY (b) IVT 132 h forecast error (shaded) and ECMO IVT verification contours (black) for 12 UTC 12 March 2020.

Forecast errors for the second IVT are associated with a developing trough moving northeast from the dateline (not shown here, but see Appendix D) which is under-predicted along with its accompanying IVT. The errors grow from 60 h onward in both the CTRL and DENY forecasts, with the CTRL errors growing faster and dominating the MAE statistic from 72-156 h. IVT-related MAEs, and MAEs for all model forecast variables, are a consequence of the GFS IVT forecast persisting the eastward IVT path (Fig. 45), i. e., following its predecessor, until landfall at 132 h (12 UTC 12 March) vice the IVT verification path that continues to follow the strengthening and northward moving cyclone (Fig. 46a, b), resulting in an eventual landfall over the Aleutians (see Appendix D for additional details). Consecutive CTRL and DENY forecasts

valid at 12 UTC 12 March (12 h before landfall) from IOPs 13-17 (Table 2) show that the above forecast scenario continues, but to a lesser extent, for IOPs 14-15 (8-9 March) until the error growth is reduced by the assimilation of later IOP data and the shortened lead time to landfall for IOPs 16-17, at which time the CTRL has reduced error compared to DENY.

IOP	Forecast hour	Initial date (00 UTC)	CTRL MAE (kg/m/s)	DENY MAE (kg/m/s)
13	132	7 March	88.6	76.3
14	108	8 March	81.0	75.6
15	84	9 March	60.3	60.1
16	60	10 March	30.4	34.1
17	36	11 March	25.2	26.0

Table 2. Domain-wide IVT MAE for CTRL and DENY forecasts from consecutive IOPs 13-17, all valid at 12 UTC 12 March. The forecast hour and initial dates are also tabulated. The verification domain is #4 (Table 1).

During IOP-13, observations were not deployed for prospective landfalling IVTs over Baja and southern California. However, the IVT-13 forecast includes multiple significant impacting events from the merged cyclone-IVT system over this area (Table 3). Over 12-48 h, an IVT makes landfall over south Baja and is well forecast by both CTRL and DENY experiments through 48 h. Over the forecast range 60-84 h and verifying over 12 UTC 9 March to 12 UTC 10 March, the system begins to be influenced by the offshore cyclone and the IVT progresses northward and reaches the US-Mexico border. At this time, both CTRL and DENY forecasts are too strong and displaced to the north, but the CTRL MAE is reduced. During the 96-108 h forecast period, a second IVT maximum approaches Baja: both the CTRL and DENY IVTs are displaced further to the south, with the CTRL more so. In the last event, from 132-144 h of the forecast, the IVT splits into two branches, with the western branch extending northward to eastern Arizona and the second branch heading northeast into Texas and Oklahoma. Over this 12 h period, the CTRL IVT MAE is much larger than the DENY (e. g., Fig. 44a, b) and reflects the southern displacement of the IVT and the weak extension into the southern USA. Later on, major precipitation occurs in both of these areas over 13-14 March.

Event	Description	Forecast event range (h)	Forecast hour	Verification date	CTRL MAE (kg/m/s)	DENY MAE (kg/m/s)
1	Initial landfall	12-48	48	00 UTC 9 March	48.4	45.7
2	Cyclone influence	60-84	84	12 UTC 10 March	57.6	68.5
3	Second IVT maximum landfall	96-108	108	12 UTC 11 March	66.4	53.7
4	Baja IVT splits into two streams	132-144	144	00 UTC 13 March	94.2	71.4

Table 3. Summary of significant system events for the Baja California IVT as covered by the IOP-13 12-168 h forecast. A brief description of the event, the covering forecast range, verification date and CTRL and DENY MAEs are tabulated. The verification domain is #8 (Table 1, Fig. 1).

(iv) *Impact Summary*

Forecast-averaged IVT impacts for all IOPs (Fig. 32) were sorted into 5 categories and ranked from 1-5, where category 1 contains the largest (3) negative IOP impacts (Fig. 47a), category 5 the largest (3) positive impacts, category 3 contains the middle (5) IOP forecast impacts and categories 2 and 4 are intermediate impacts (3 each). Likewise, forecast-averaged impacts for the model forecast variables were similarly sorted and categorized after vertical averaging for each IOP: geopotential height and wind speed were averaged over the 200-925 hPa layer and moisture was averaged over the 700-850 hPa layer. IOPs with the highest/lowest IVT impact (IOP-2, IOP-13 respectively) are also in the highest/lowest categories for the model forecast variables (Fig. 47b). IOP-10 (21 February) also has consistent degradation for IVT and the model forecast variables. IOPs with lower magnitude IVT impacts have lower impacts from forecast variables (e.g., IOP-4) or mixed impact across the forecast variables. For example, IOP-7 has positive impact for IVT and geopotential height but wind speed has a lesser positive impact and a negative moisture impact. Category 3 IOPs for IVT generally have model variable impacts in the same category. Again, IOP-8 is an unusual case with category 5 impact for IVT, wind speed and moisture but a negative impact for geopotential height as noted above.

(a)

Rank	Key	#IOPs
1		3
2		3
3		5
4		3
5		3

(b)

IOP	Date	IVT	Z	WSPD	SPCH
1	00Z24JAN2020	-1.45	0.26	0.78	1.07
2	00Z29JAN2020	6.82	6.12	7.33	5.88
3	00Z31JAN2020	-3.65	-7.05	-4.70	-1.23
4	00Z04FEB2020	1.64	3.24	2.91	2.47
5	00Z05FEB2020	-1.86	1.86	-0.47	0.47
6	00Z06FEB2020	-1.06	3.90	0.31	2.39
7	00Z14FEB2020	2.15	5.05	1.22	-2.25
8	00Z15FEB2020	3.40	-4.12	4.33	5.47
9	00Z16FEB2020	-2.01	-2.36	-0.32	0.20
10	00Z21FEB2020	-3.93	-9.23	-4.03	-4.14
11	00Z24FEB2020	1.87	1.33	1.51	-0.15
12	00Z02MAR2020	2.05	9.82	4.19	2.50
13	00Z07MAR2020	-11.02	-18.16	-11.39	-5.87
14	00Z08MAR2020	-2.09	-1.29	0.98	3.15
15	00Z09MAR2020	1.19	0.08	0.75	-0.37
16	00Z10MAR2020	1.24	-1.13	1.11	2.75
17	00Z11MAR2020	-1.32	-0.20	-0.51	1.58

Fig. 47. Ranked impact (I) categories (a) for forecast-averaged IVT and forecast- and vertically-averaged geopotential height, wind speed and specific humidity impacts (b). See text for processing details.

3) Precipitation impacts

Positive precipitation forecast impacts are probably the most important desired result from the AR Program. Model forecast precipitation itself, however, is a complex function of the forecast divergent wind, available moisture and thermodynamic vertical stability and the various physical parameterizations in the model itself. The IOPs were designed to improve precipitation forecasts over the WEST, PNNC and SCAN domains. Generally, IOPs 1-9 and 11-14 (Table 1) were focused on future events in the PNNC domain and IOPs 10, 12 and 14-17 were focused on the SCAN domain. However, some major events in these domains were not addressed by any IOP observations (e. g., 14-16 March in the PNNC domain and 25 January – 16 February in the SCAN domain), so that any impacts (positive or negative) at these dates and in these domains would be indirect due to the data assimilation system carrying prior IOP observations forward in time. After presenting the forecast precipitation statistics used here, the time series of observed and model forecast precipitation over the WEST, PNNC and SCAN domains are described, and regional scores for the entire OC together with direct impact statistics for relevant IOPs for the PNNC and SCAN domains are presented.

(i) Forecast precipitation statistics

Both precipitation magnitude, duration and timing are critical criteria for measuring forecast success. Operational standard “threat” and “bias” scores (Wilks 2006; Jolliffe and Stephenson 2003) include contributions from each of these criteria and are presented first. Additional statistical verification for precipitation forecasts mirrors the procedures for IVT events and model variables as already presented in Sections 1-2 above. Statistics capture improvements/degradations for CTRL/DENY “full field” (FF) quantities such as domain-wide maximum and domain-averaged precipitation values relative to their CCPA counterparts as well as domain-wide maximum and minimum values and standard deviations of gridded “difference fields” (DF, CTRL-CCPA and DENY-CCPA). Precipitation statistics were generated for the WEST, PNNC and SCAN domains with verifications at both 00 and 12 UTC for the entire AR2020 OC (00 UTC 24 January to 12 UTC 11 March, a total of 96 cases). In addition, these statistics were then sampled for the critical forecasts of each IOP (Table B1).

Basic “Full-Field” (FF) verified quantities are the maximum and domain-averaged precipitation for each model on its Gaussian grid within each domain, $FFpmax_{CTRL}$, $FFpmax_{DENY}$, $FFpavg_{CTRL}$, and $FFpavg_{DENY}$, respectively, and the corresponding observed (CCPA) maximum and average amounts on its grid, $FFpmax_{CCPA}$ and $FFpavg_{CCPA}$. All quantities are verified for 24-120 h forecasts against 24 h accumulated CCPA ending at either 00 or 12 UTC on the verifying date as appropriate. The first verification date is 00 UTC 25 January for the 24 h forecast and the final date is 12 UTC 16 March for the 120 h forecast. The $FFpavg$ and $FFpmax$ derived statistics (S) are the absolute values of differences, e. g.,

$$S_{CTRL} = |FFpmax_{CTRL} - FFpmax_{CCPA}|, \quad (3)$$

And are used in (2) to calculate the error improvement (I) for both maximum and domain-averaged precipitation; again, positive values of I indicate that CTRL maximum/average precipitation values are closer to observed values than for DENY.

The basic “Difference Field” (DF) statistics are calculated after interpolating the FF model precipitation output to the CCPA grid and taking the model-minus-observed difference over the domain. The minimum, maximum and standard deviation of the difference ($DFpmin$, $DFpmax$ and $DFpstd$ respectively) for each model run are then compared and the improvement calculated as, for example:

$$I_{DFpmin} = |DFpmin_{DENY} - DFpmin_{CTRL}| / |DFpmin_{DENY}|. \quad (4)$$

Improvements of CTRL over DENY for these statistics indicate a better domain-wide geographical distribution of the DF wherever the CCPA is defined.

Forecast statistics are calculated for the 24-120 h range and each domain (WEST, PNNC and SCAN) by averaging over all valid dates for each forecast hour for indirect statistics and each IOP for direct statistics. Since these statistics are noisy, resulting from both naturally noisy precipitation distributions and a very small number of verifications, results are best described categorically as improvements ($I > I_{thr} > 0$), degradations ($I < -I_{thr} < 0$), where $I_{thr} = 0.5\%$ and unchanged ($-I_{thr} \leq I \leq I_{thr}$). To prevent unrealistic calculations from distorting the statistics, a filter is applied to all calculations; further details on calculation of the precipitation statistics are given in Appendix E.

(ii) Time series of observed and model forecast precipitation

The time series of 24, 72 and 120 h CTRL forecast precipitation, averaged over the PNNC domain, shows good correspondence with observed values over the OC (Fig. 48a), but with a 10% increase in mean value at 120 h compared to 24 h, while the 72 h and 24 h means are equivalent. At individual verification times, however, there are patches of strong over-prediction (00 UTC 26 January to 00 UTC 27 January, 00 UTC 16 February to 00 UTC 17 February, 00 UTC 2 March to 00 UTC 3 March and 00 UTC 9 March to 12 UTC 9 March) and other patches of notable under-prediction (12 UTC 8 February to 12 UTC 9 February and 00 UTC 14 March to 12 UTC 16 March). Over-prediction is worse at 120 h than shorter forecasts and under prediction is without trend across all forecast hours.

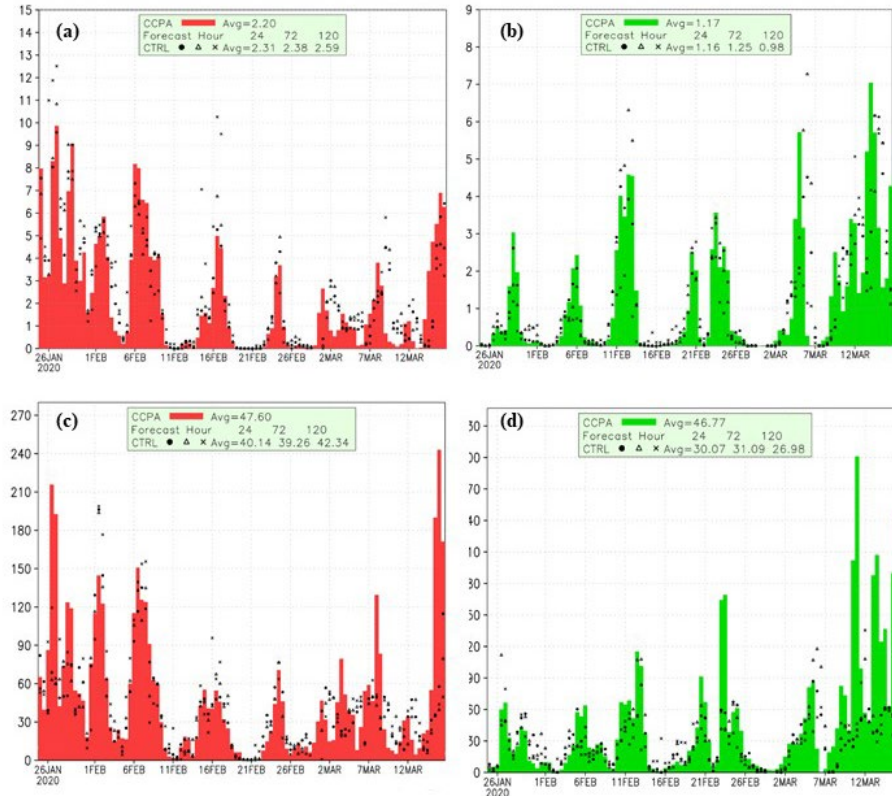


Fig. 48. As in Fig. 2 over the PNNC (a, c) and SCAN (b, d) domains for area-average and domain maximum precipitation, but with added CTRL domain-average and maximum forecast precipitation for 24 h (dot), 72 h (triangle) and 120 h (X) at both 00 and 12 UTC for the AR2020 OC.

The CTRL forecast average precipitation over the SCAN domain (Fig. 48b), however, shows a 16% decreasing trend compared to the CCPA from 24 h to 120 h. The number and magnitude of the over-predictions is somewhat reduced compared to the PNNC domain, but there are over-prediction periods from 00 UTC 12 February to 00 UTC 13 February, 12 UTC 5 March to 12 UTC 6 March, and various times from 12 UTC 12 March to 12 UTC 15 March for some forecast hours but not others. Due to the opposite trends in time- and domain-averaged precipitation in the PNNC and SCAN domains, the WEST domain, which encompasses both sub-domains, exhibits fewer positive outliers and the forecast trends are hidden.

For PNNC domain maximum precipitation, the CTRL has low biases of -15.7%, -17.5% and -11.1% for 24, 72 and 120 h forecasts respectively over the PNNC domain (Fig. 48c). Undoubtedly, some of this bias is model-resolution dependent, but there is at least one period, 00 UTC 1 February to 00 UTC 2 February when the model maxima for all forecast hours exceeds the observed CCPA and several others when at least one forecast hour does also (e. g., 00 UTC 7 February to 00 UTC 8 February, 00 UTC 14 February to 12 UTC 17 February and 00 UTC 9 March to 12 UTC 13 March).

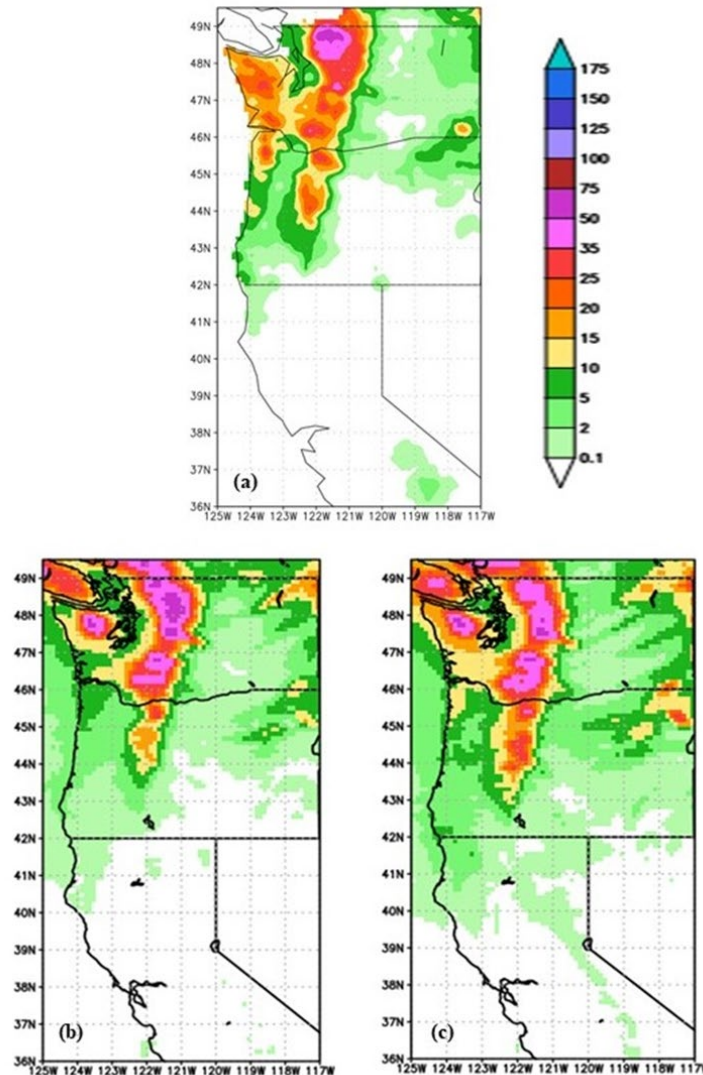


Fig. 49. 24 hour CCPA accumulated observed precipitation (mm) ending at 12 UTC 24 February (a) and CTRL (b) and DENY (c) 60-84 h forecast precipitation (mm) valid on the same date.

The maximum precipitation over the PNNC domain for the event of 22-24 February, however, is well forecast by the CTRL (Fig. 49b). Early in this period a cyclone forms northwest of Vancouver Island with an IVT to the southeast and strong onshore moisture flux over northwest Washington State. The 60-84 h accumulated observed and forecast precipitation, valid 12 UTC 24 February, has maxima of 71 (CCPA), 64 (CTRL) and 56 mm (DENY), so that the CTRL maximum is much closer to observed than the DENY and has an improvement of 53%; moreover, the broad area of maximum precipitation in north Washington State is improved in the CTRL relative to DENY. Difference fields of DENY/CTRL with the CCPA (Fig. 50a, b) also show that the DENY is under forecast by 10-20 mm in the maximum precipitation area near the US-Canadian border, while the CTRL, while still under forecast, is less so and is an improvement of 4.5% in the ID_{Fpstd} statistic.

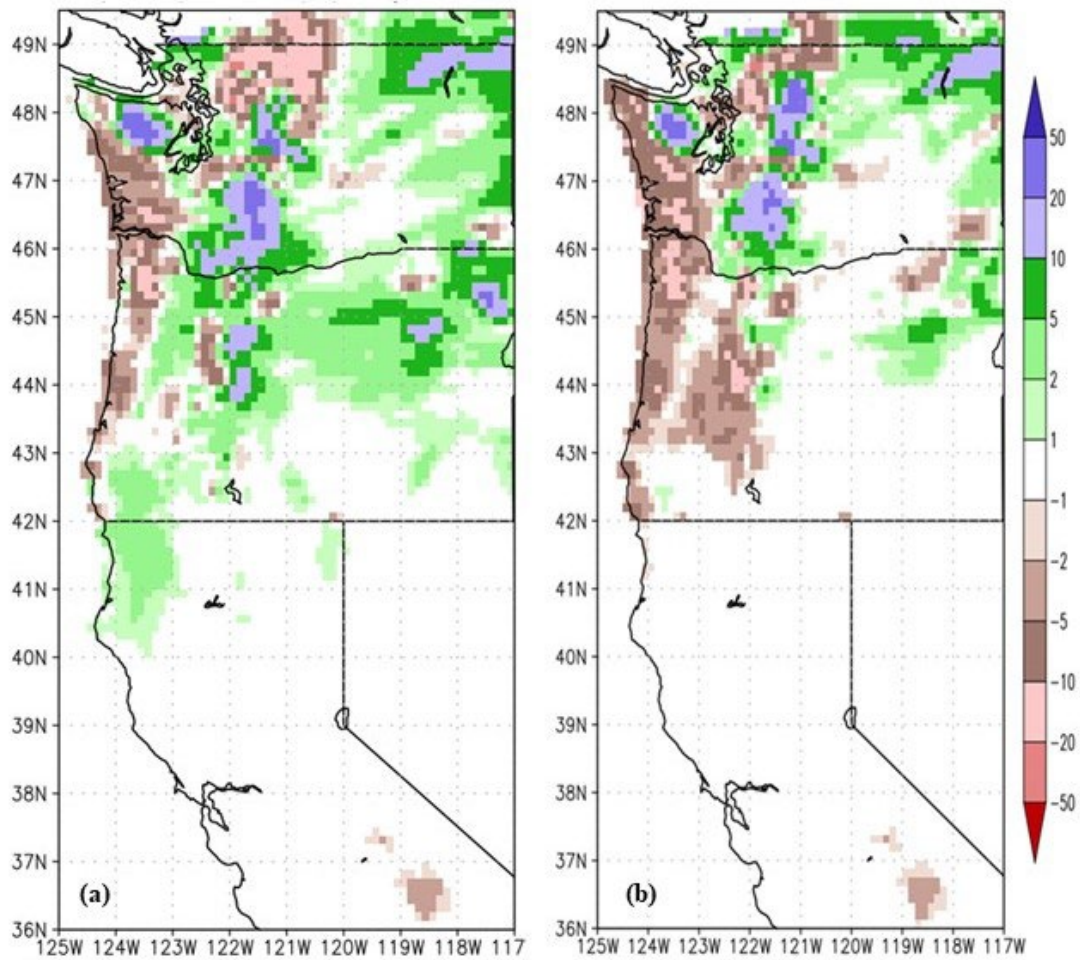


Fig. 50. Difference of the DENY (a) and CTRL (b) 60-84 h forecast precipitation (mm) from the CCPA observed precipitation valid on 12 UTC 24 February.

Over the SCAN domain, the maximum precipitation amounts (Fig. 48d) exceed the observed in fewer cases than for the PNNC, but some exceptions (i.e., false alarms) occur for low precipitation periods, from 00 UTC 30 January to 00 UTC 4 February, for example. The CTRL maximum precipitation biases are much lower for the SCAN domain than for PNNC: -35.7%, -33.5% and -42.3% for 24, 72 and 120 h forecasts respectively.

(iii) Regional scores

For continental scales, precipitation impacts are insignificant (not shown) just as for the standard model variables described earlier. However, regional statistics over the U.S. West Coast, viz. the operational threat and bias scores for 00 UTC initializations valid at 12 UTC (Fig. 51a, b) over a domain approximately equal to the WEST domain show some positive impact for the threat score, particularly for moderate rainfall amounts (10 mm/day) at 132-144 h and for heavy rainfall (35 mm) at 96-120 h. Furthermore, time-averaged 108-132 h CTRL threat scores for all precipitation amounts (Fig. 52b) are improved over DENY at the 99% statistically significant level for two precipitation categories.

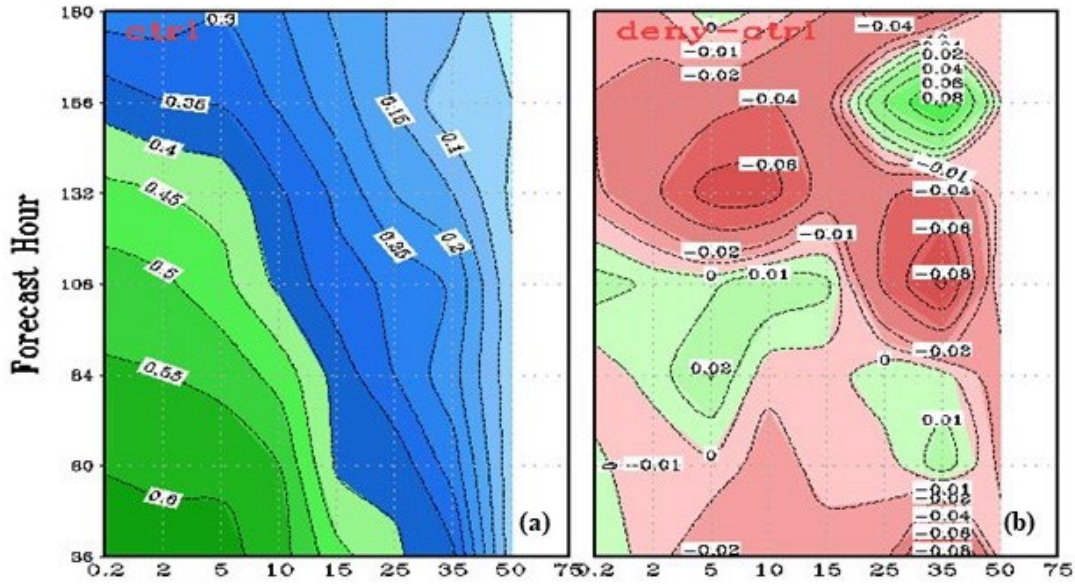


Fig. 51. CTRL threat score for the experimental period (a) and DENY-CTRL difference (b) for f036-f180 over the U. S. West Coast region (approximately 32-49.5N, 115-125W). The score is for 0.2-75 mm/day precipitation thresholds. Positive impact (b) is red and negative impact is green.

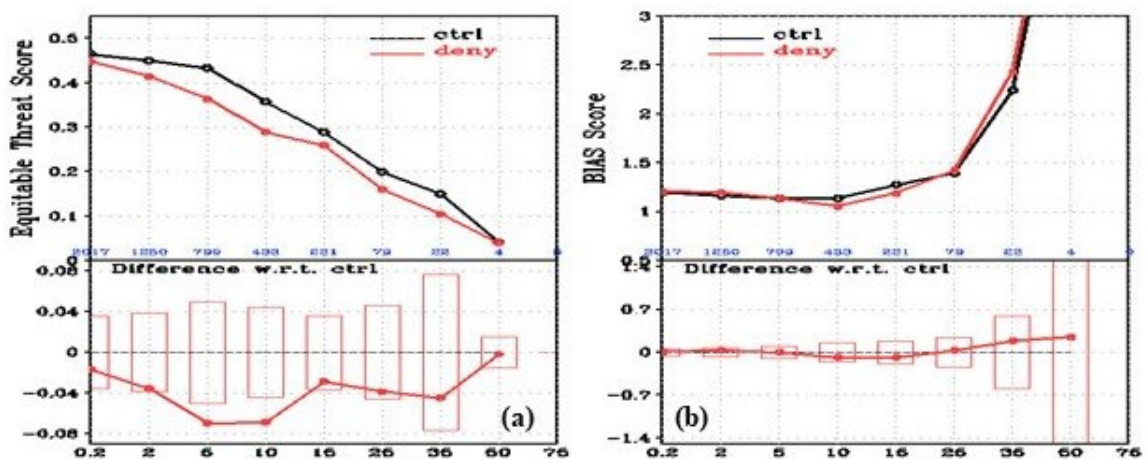


Fig. 52. 108-132 h forecast CTRL (black) and DENY (red) threat scores (a, top) and bias scores (b, top) and DENY-CTRL differences (bottom) for the AR2020 experimental period over the U. S. West Coast region (approximately 32-49.5N, 115-125W). Categories on the abscissa are 24-h accumulations (mm). Differences outside the vertical boxes indicate statistical significance at the 99% level.

Forecast improvement statistics of CTRL over DENY (e. g., Eq. 4), averaged over all 00 and 12 UTC cases for the WEST domain (Fig. 53a), show generally positive impact of 1-4% in the 24-60 h range, with the domain-averaged precipitation improved by almost 4% at 60 h, but mixed impact for the domain-wide maximum precipitation. At 72-84 h, difference statistics are preponderantly negative, but the domain-wide FF maximum and average precipitation statistics I_{FFpmax} and I_{FFpavg} , are uniformly positive. For longer forecasts (96-120 h), most statistics indicate positive impact, consistent with the regional threat scores (Figs. 51, 52). Over all forecast ranges, there is a 5-15% majority of improved cases (Fig. 53b) for most FF and DF statistics, except for $IDFpmin$. The latter exception could arise from a positive model precipitation bias.

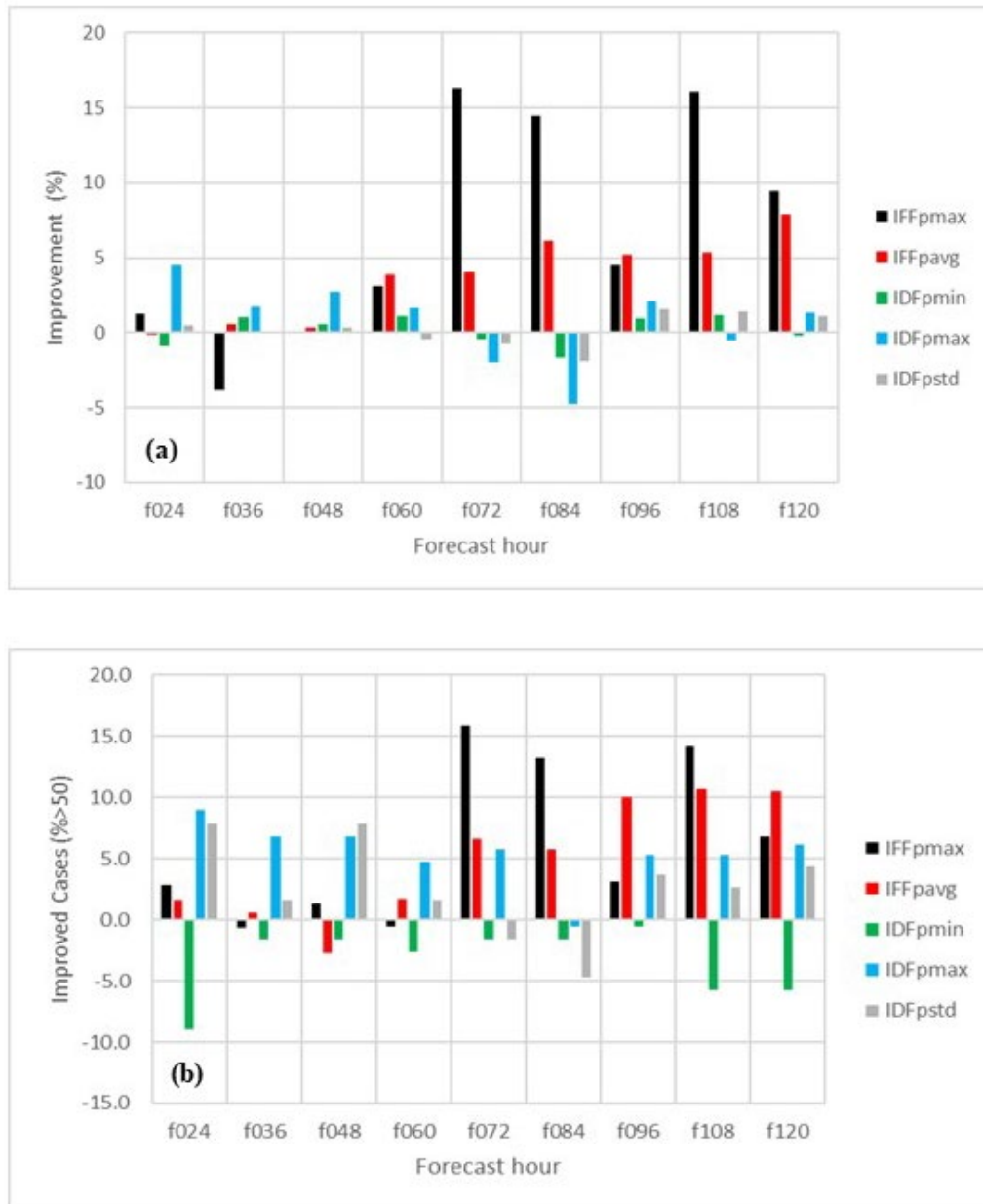


Fig. 53. Average improvement (a, %) of CTRL over DENY and percent improved cases (b, %> 50) for 96 twice-daily (00 and 12 UTC) forecast cases and 24-120 h forecasts over the WEST domain (Table 1). Forecasts are initialized from 00 UTC 24 January to 12 UTC 11 March. Statistics are as described in the text.

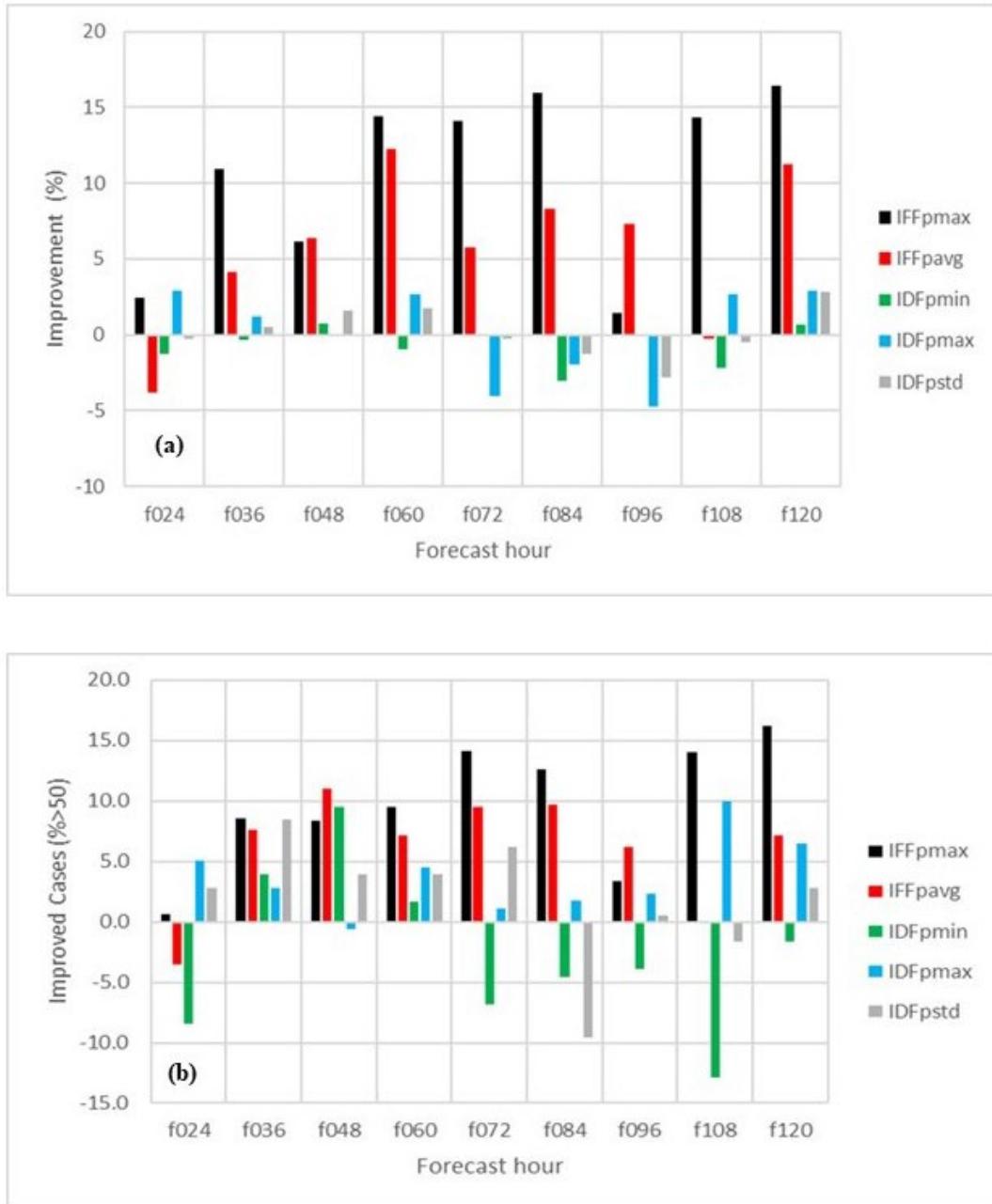


Fig. 54. As in Fig. 53, except for the PNNC domain.

PNNC domain statistics are mixed for 24 h forecasts (Fig. 54a) and show mostly small positive impacts of <5% for DF quantities for 36-60 h forecasts, generally negative impact of at most -5% at 72-96 h and mostly positive impact at 108-120 h. It is, however, notable that FF statistics (IFF_{pmax} and IFF_{pavg}) continue positive at approximately 5-15% throughout 36-120 h, indicating that the dropsonde data result in domain-wide maximum and average precipitation rates closer to CCPA observations in the PNNC domain, as illustrated by Fig. 50. Moreover, corresponding case majorities (Fig. 54b) are approximately 7-15% for FF quantities and generally positive for most DF quantities, thereby indicating overall improved geographical precipitation distribution from the CTRL forecasts relative to DENY.

The above case-averaged improvement statistics, after averaging over all forecast hours and all statistics, are positive for all domains, with a maximum improvement of 3.41% for the PNNC domain (Table 3) and the largest plurality of positive cases (3.8%). Improvements for the PNNC domain are most encouraging for the maximum forecast value (I_{FFpmax} , 11.7%) and the domain-average value (I_{FFpavg} , 5.7%), each of which also has a strong accompanying plurality of positive impacts. While the SCAN domain has smaller impacts overall, the individual statistics indicate positive impacts and positive plurality in all but one quantity (I_{DFpmin}). More details on impacts, including for the SCAN domain and a summary of all cases for the I_{DFpstd} statistic, are given in Appendix E.

Improvement statistic	WEST	PNNC	SCAN
I_{FFpmax}	6.83 (6.2)	11.72 (9.7)	5.9 (5.2)
I_{FFpavg}	3.71 (5.0)	5.71 (6.1)	4.22 (0.0)
I_{DFpmin}	0.18 (-3.3)	-0.72 (-2.6)	-0.54 (-10.7)
I_{DFpmax}	0.76 (5.5)	0.17 (3.7)	2.19 (5.3)
I_{DFpstd}	0.21 (2.6)	0.19 (1.9)	0.18 (0.1)
Average	2.34 (3.2)	3.41 (3.8)	2.39 (-0.02)

Table 4. Improvement statistics of CTRL over DENY forecast precipitation averaged over 96 cases and forecast hours 24-120 for the WEST domain and the two sub-domains, PNNC and SCAN (Table 1). The number in parentheses is the percent of positive impact cases, expressed as a majority (i. e., > 0 means more than 50% improved cases and vice versa for negative values). The last line is the average value of all statistics for each domain.

(iv) Direct impacts from IOPs only

For each IOP (except IOP-11, which has no CCPA verification over western Canada), precipitation statistics were selected for critical forecast hours over the more relevant, more local, PNNC and SCAN verification domains (Table B1). Consistent with operational precipitation verification at 12 UTC, verification of 00 UTC initializations is done for forecast ranges 12-36 h, 36-60 h, etc. Overall, results are mixed for each forecast hour (Fig. 55). Precipitation forecasts for 96 h are improved in 4 of 5 statistics. OC-averaged I_{DFpmax} and I_{DFpstd} are positive for longer forecasts, implying an improved geographical precipitation distribution. Short-range forecasts (24-36 h) are mostly not improved, while 48-60 h forecasts have improved I_{DFpstd} . At 72-84 h, forecasts are largely degraded or unchanged across multiple statistics. These statistics are consistent with error reductions for the mass, wind and specific humidity fields (Figs. 36-39). A breakdown of these statistics by IOP and with percent performance changes is given in Appendix E.

Forecast Hour	I _{FFpmax}	I _{FFpavg}	I _{DFpmin}	I _{DFpmax}	I _{DFpstd}	IOPs	Valid Dates	Statistics Improved	Statistics Degraded	Statistics Unchanged
f024	Red	Green	Red	Yellow	Red	9	25	1	3	1
f036	Red	Yellow	Yellow	Green	Yellow	12	33	1	1	3
f048	Red	Red	Green	Red	Green	11	25	2	3	0
f060	Red	Red	Yellow	Green	Green	11	27	2	2	1
f072	Red	Green	Red	Red	Yellow	9	18	1	3	1
f084	Yellow	Green	Red	Red	Yellow	9	21	1	2	2
f096	Green	Green	Green	Red	Green	7	13	4	1	0
f108	Red	Yellow	Red	Green	Green	6	12	2	2	1
f120	Red	Green	Red	Green	Green	3	6	3	2	0
f132	Red	Yellow	Green	Green	Green	3	6	3	1	1

Fig. 55. Summary of precipitation improvement statistics for 24-132 h, including all IOPs in the AR2020 OC that verify during a period of high impact precipitation. Improved (green), degraded (red) and unchanged (yellow) impacts are indicated for each forecast hour. The number of IOPs for each forecast hour, the total number of valid dates and the number of statistics improved, degraded and unchanged are also tabulated.

6. Summary and discussion

Aircraft-deployed wind and thermodynamic soundings, ingested into the NCEP GDAS during the AR2020 OC, provided short-term forecast error reductions for the analysis and short-term (12-24 h) basic forecast variables (wind, moisture, mean sea-level pressure and geopotential height) but not for IVT maxima (analysis and 24 h only are categorized as improvements, Fig. 30). Within the AR phenomenon, the IVT geographical extent and strength are more difficult features to verify since they are localized and intermittent; a larger sample size is most likely required to demonstrate impact. Although regional precipitation forecast improvements are demonstrated for the entire OC for the WEST and PNNC domains (Figs. 51-54, Table 4), more local rainfall improvements over IVT landfalling areas are, however, not clearly shown, although the PNNC case verifying at 12 UTC 24 February is one example of an improved forecast (Figs. 49-50).

Additional in-situ moisture information provided by the AR2020 dropsondes appears to have a positive forecast impact as it is a unique and accurate measurement over most oceanic areas and of sufficient density to characterize the ARs. Improved IOP-averaged CTRL moisture verification at 120 h and beyond is notable for this OC (Fig. 39). It is likely that this improved moisture transport is associated with any precipitation improvements in the SCAN domain and at longer forecast hours (120-168 h) over the WEST and PNNC domains.

Episodic dropsonde deployments impact larger areas both upstream and downstream of the verification area through the GDAS cycling process as time proceeds. Observations for a particular AR event may impact subsequent AR events as they enter the verification area beyond 36-48 hours. The nature of this impact is generally unpredictable, but for very limited observation coverage it is more likely to be negative or neutral than positive. Nevertheless, some positive regional impacts for 24-120 h precipitation forecasts have been shown here (Table 4) and for 108-168 h elsewhere (Wu et al. 2021); they coincide with improved wind speed and moisture statistics over the entire OC (Figs. 38-39) and may be responsible for the small average IVT MAE improvements in the 120-132 forecasts (Fig. 30) and increased numbers of improved IOP cases over this forecast period (Fig. 34c).

Due to small sample sizes, statistics averaged over the 17 IOPs can be dominated (in either direction) by a single case. For example, the improvement at 144 h for wind speed and moisture (Figs. 38, 39) comes primarily from IOP-14 (00 UTC 8 March deployment, Figs. 40c, d), which

is large compared to all other impacts (not shown). Indeed, there is considerable impact variability in the complete forecast sample for forecast length, IOP and model variable. Furthermore, small forecast differences in the short range can occasionally lead to major subsequent forecast errors (e.g., see Appendix D, section 3 and Fig. D10). While the mechanisms described by Hodyss and Majumdar (2007) and Ancel et al. (2018), related to spreading of model-generated noise cannot be ruled out, tracking of the forecast differences between CTRL, DENY and the verifying analyses has revealed error mechanisms that appear physically realistic. Given that the cycled data assimilation spreads the inserted dropsonde information globally over time, occasional large forecast differences are bound to occur.

This study also documents some heightened PBL moisture variability in the GFS/GDAS system relative to that in the ECMWF forecast system. A more substantial study of the low-level moisture distribution, relative to the dropsonde data, may prove useful in diagnosing this issue.

The AR Recon program has evolved from a field demonstration to a real-time operational capability as documented in the National Winter Season Operations Plan by the Office of Federal Coordinator of Meteorology (OFCM, 2020), thereby providing more opportunities to thoroughly investigate the impact of aircraft observations on NCEP's operational global model analysis and forecasts. AR Recon in 2021 consisted of multiple aircraft and several sequential (multi-day) IOPs, giving unprecedented coverage of ARs. In addition, NCEP made significant advancements to the operational GFS and GDAS in 2021 (GFSv16, Yang et al. 2021, Kleist et al. 2021). For the first time, real-time data denial experiments were conducted using GFSv16. Results from those experiments will be documented in a sequel to this manuscript, along with focused examination of precipitation forecast improvements at local watershed levels where the impacts are found to be more pronounced.

Acknowledgments.

This work is supported by generous funding made available from NOAA Office of Marine and Aviation Observations (OMAO). The authors are grateful to the feedback from the Modeling and Data Assimilation Steering Committee of the AR Recon Program, and to the internal and external reviewers of this manuscript.

Data Availability Statement.

All model data, observations and statistical results are available through the corresponding author at the NOAA National Centers for Environmental Prediction, Environmental Modeling Center.

APPENDIX A

AR2020 Intensive Observing Periods

Table A1 provides some details on the 17 AR2020 IOPs, including dates, participating aircraft, number of dropsondes (and failures) from each of the G-IV and C-130 aircraft and the number of observations assimilated in real time by the GDAS for the 00 UTC cycle and the surrounding cycles (18 and 06 UTC). In some IOPs, (#s 6, 7, 15 and 16), not all successful dropsondes were transmitted to NCEP and other centers. These missing data have not been recovered.

IOP #	Date 2020	AF C-130	AF C-130	NOAA G-IV	GDAS Assimilated IOP Dropsondes	GDAS Cycle (18/00/06 UTC)
		No. of	dropsondes	(failed)		
1	00 UTC 24 Jan	25 (0)	14 (1)		38	1/37/0
2	00 UTC 29 Jan	26 (1)	2 (0)		27	(2/21/4)
3	00 UTC 31 Jan		25 (1)		24	0/24/0
4	00 UTC 4 Feb	19 (4)	25 (0)	35 (5)	70	0/70/0
5	00 UTC 5 Feb			34 (4)	30	0/30/0
6	00 UTC 6 Feb	25 (3)	26 (1)	35 (5)	59	0/59/0
7	00 UTC 14 Feb			30 (0)	27	0/27/0
8	00 UTC 15 Feb	26 (1)	31 (6)	35 (5)	80	0/78/2
9	00 UTC 16 Feb			30 (0)	30	0/30/0
10	00 UTC 21 Feb			30 (0)	30	1/29/0
11	00 UTC 24 Feb	29 (3)		32 (2)	56	1/55/0
12	00 UTC 2 Mar	27 (5)	25 (1)	30 (0)	76	0/74/2
13	00 UTC 7 Mar			31 (1)	30	0/30/0
14	00 UTC 8 Mar		28 (3)	31 (1)	55	0/54/1
15	00 UTC 9 Mar		23 (3)	29 (0)	29	0/29/0
16	00 UTC 10 Mar		30 (0)	27 (0)	43	0/39/4
17	00 UTC 11 Mar		2 (0)		2	0/2/0

Table A1. Summary of the AR2020 IOPs. Columns identify the serial number of each IOP, the central time for the aircraft observations, the number of dropsondes (failed number) and the number of sondes included in the 18, 00 and 06 UTC GDAS cycles.

APPENDIX B

IOP Critical Forecast Ranges

IOP	Observation Date (00 UTC)	Precipitation Date (beginning)	Precipitation Date (end)	Critical Forecast Range (h)	Domain
1	24 Jan.	24 Jan. 00 UTC	29 Jan. 00 UTC	24,...120	PNNC
2	29 Jan.	29 Jan. 00 UTC	30 Jan. 12 UTC	24,...108	PNNC
3	31 Jan.	31 Jan. 12 UTC	1 Feb. 12 UTC	24,...60	PNNC
4	4 Feb.	5 Feb. 00 UTC	9 Feb. 00 UTC	24,...120	PNNC
5	5 Feb.	5 Feb. 12 UTC	9 Feb. 00 UTC	24,...96	PNNC
6	6 Feb.	5 Feb. 12 UTC	9 Feb. 00 UTC	24,...72	PNNC
7	14 Feb.	15 Feb. 00 UTC	17 Feb. 00UTC	24,...72	PNNC
8	15 Feb.	16 Feb. 00 UTC	17 Feb. 00UTC	24,...48	PNNC
9	16 Feb.	17 Feb. 00 UTC	17 Feb. 00UTC	24	PNNC
10	21 Feb.	22 Feb. 12 UTC	23 Feb. 12 UTC	36,...60	SCAN
11	24 Feb.	26 Feb. 00 UTC	26 Feb. 00 UTC	-	-
12	2 Mar.	2 Mar. 12 UTC	5 Mar. 12 UTC	36,...84	SCAN
13	7 Mar.	9 Mar. 00 UTC	15 Mar. 00 UTC	48,...168	SCAN
14	8 Mar.	9 Mar. 00 UTC	15 Mar. 00 UTC	24,...168	SCAN
15	9 Mar.	10 Mar. 00 UTC	15 Mar. 00 UTC	24,...144	SCAN
16	10 Mar.	11 Mar. 00 UTC	15 Mar. 00 UTC	24,...120	SCAN
17	11 Mar.	12 Mar. 00 UTC	15 Mar. 00 UTC	24,...96	SCAN

Table B1. Summary of the critical forecasts for land-based precipitation associated with each IOP. IOPs are numbered as in Table A1 and dropsonde observations are deployed at 00 UTC. Critical forecast ranges (h), 24, 36, etc., are at 12 h intervals and are valid at the date in column 4, but do not include 12 h, since the latter does not contain all of the 24 h accumulated precipitation beginning at the IOP date. Local precipitation domains are the Pacific Northwest and Northern California (PNNC, domain 6) and the Southern California, Arizona and New Mexico (SCAN, domain 7) as defined in Table 1 (Fig. 1c). IOP #11 is not included for precipitation evaluation due to lack of available land-based precipitation.

APPENDIX C

A Case Study of Forecasts Initialized at 00 UTC 29 February 2020

The 120 h forecast impact from 00 UTC 29 February 2020 is a rare case of substantial indirect impact as defined in this paper. The CTRL initialization contains no dropsonde data and the closest previous IOP was at 00 UTC 24 February, 5 days earlier. Analysis differences at 00 UTC 29 February (Fig. C1) are of magnitude 2-4 m throughout most of the domain, with isolated minimum/maximum differences of -13 m and 9 m respectively. The impact at f120 is the largest Z500 improvement over the experimental period (Fig. 26). Forecast MAE (Fig. 27) increases linearly and almost equally with forecast hour for both CTRL and DENY for the first 96 h. After 96 h, the DENY MAE increases much faster before leveling off at 144 h.

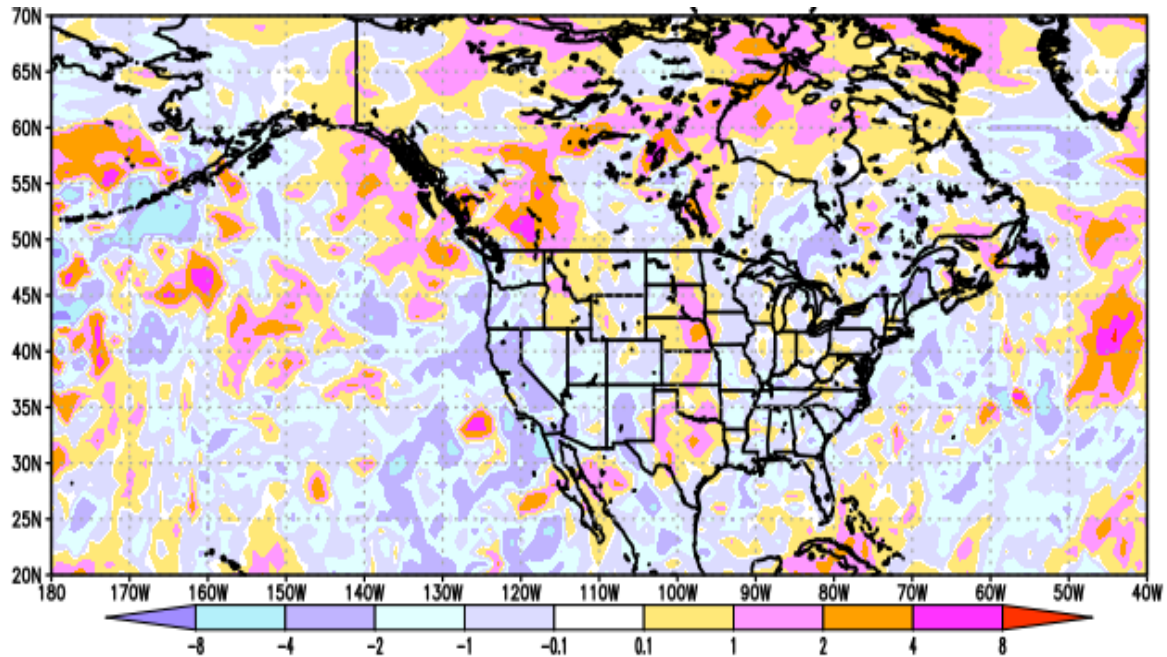


Fig. C1. Z500 analysis difference between CTRL and DENY experiments at 00 UTC 29 February. Minimum/maximum differences in the domain (20-75 N, 180-320 E) are -13.4/9.1 m, the mean difference is -0.3 m and the mean absolute difference is 1.2 m.

At 96 h, in comparison with the ECMO verifying analysis (Fig. C2a), both CTRL and DENY forecasts exhibit much the same error patterns (Figs. C2b, c). The CTRL forecast exhibits maximum positive forecast errors exceeding 100 m due to an overdeveloped ridge in southwest Canada, a progressive trough over Texas instead of a cutoff low over Arizona, New Mexico and northern Mexico, and a developing ridge with errors exceeding 150 m over the eastern US instead of a shallow trough over the Great Lakes. These error patterns are accentuated in the DENY forecast, especially as it develops a deeper trough from Texas to Minnesota (errors less than -160 m) and amplifies the ridge (errors exceeding 180 m) from the Great Lakes to New England.

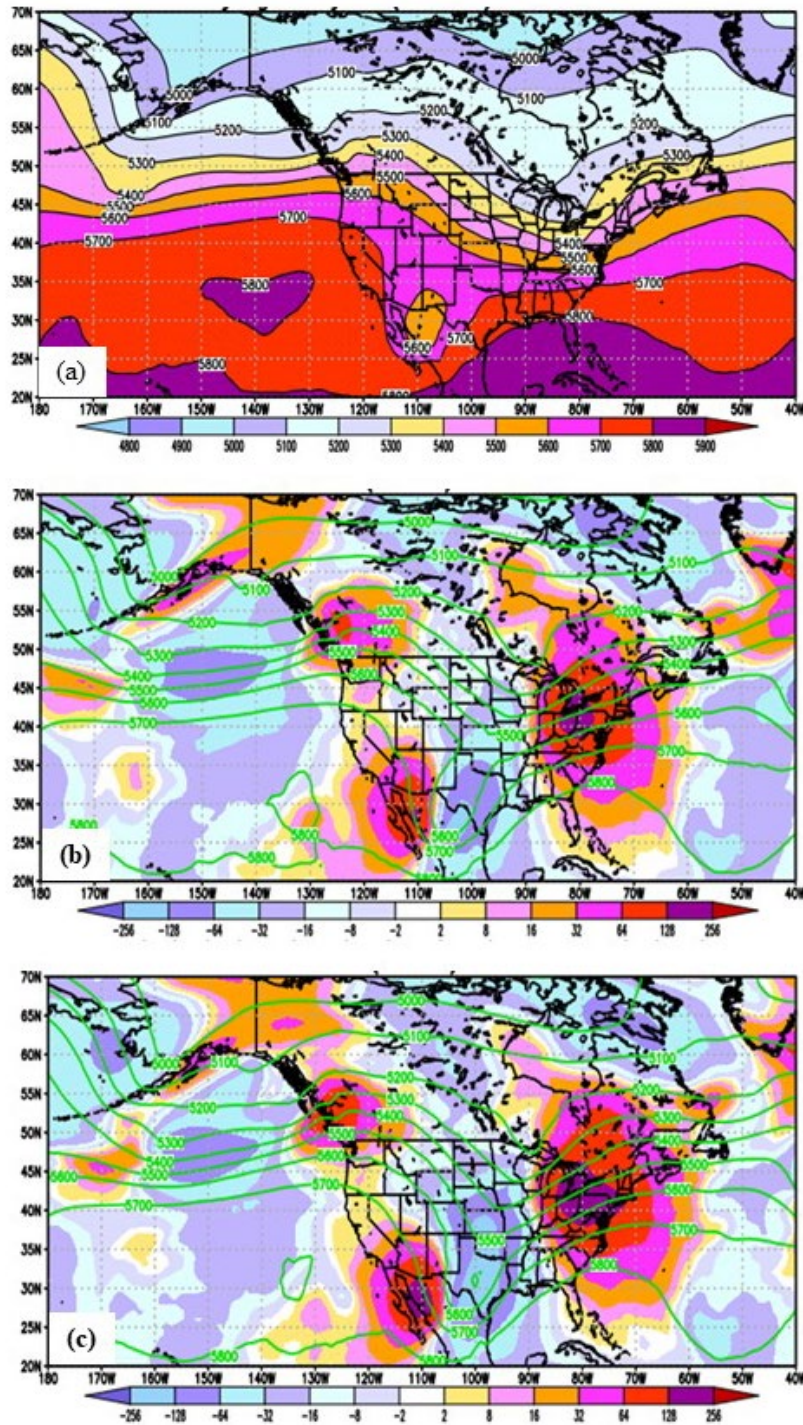


Fig. C2. Z500 ECMO verifying analysis at 00 UTC 4 March 2020 (a) and Z500 CTRL (b) and DENV (c) forecast errors at 96 h. Solid colors depict the sign and magnitude of the f096 forecast error and green contours depict the CTRL forecast.

At 120 h, the ECMO verifying analysis (Fig. C3a) shows a ridge over the western US and southwestern Canada and a trough extending southeastward from Hudson's Bay to the Canadian Maritime Provinces. The CTRL exhibits a weak trough southeast of Hudson's Bay but does not extend it sufficiently eastward as in the verification. CTRL errors exceed 270 m in this region. DENV errors exceed 390 m as it forecasts a ridge over the Maritime Provinces down through New

England. Over the Great Lakes region, CTRL predicts a weaker trough than DENY, but is still too progressive as the storm system actually develops 12-24 hours later (not shown). Even further out in the forecast, at f168 (not shown), DENY has the largest forecast errors over the Mid-Atlantic region due to the Great Lakes trough having moved through that area in advance of the verification. This is a case of fast-growing error development which produced an indirect impact on AR2020 forecasts, but did not impact the important AR region of the western US and Canada.

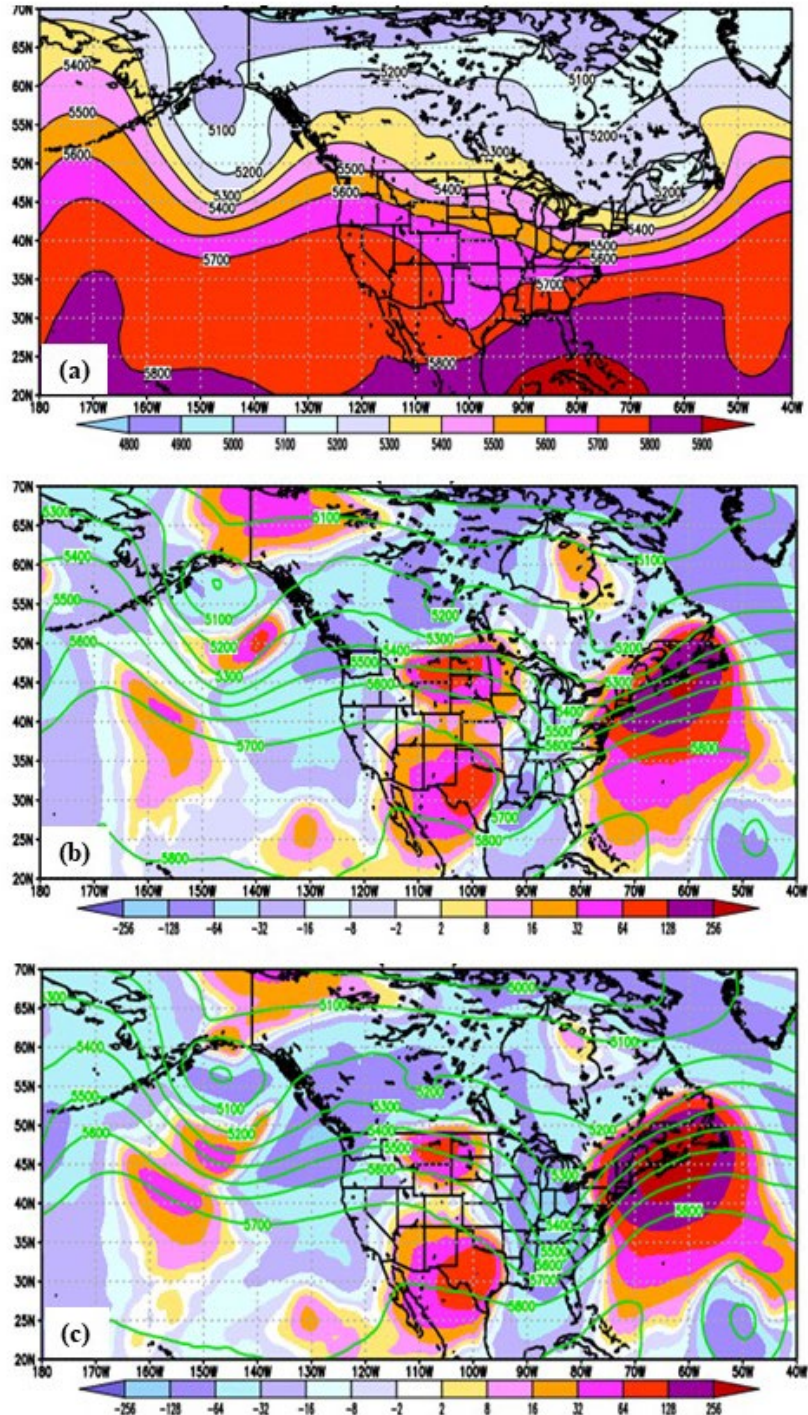


Fig. C3. As in C2, except for valid date 00 UTC 5 March 2020 and 120 h forecasts.

APPENDIX D

Impacts for Selected AR IOP Cases

The general synoptic situations, their chronologies and the associated observation have been described for each IOP in Section 4a-h. Below are more detailed descriptions of the impacts for some illustrative IOP cases: IOPs 2, 8-9 and 13.

1. IOP-2

The IOP-2 forecast covers a landfall case in BC, Canada and northwestern Washington State. It illustrates an overall positive impact due to dropsondes for the CTRL forecast initialized at 00 UTC 29 January. The AFRES C-130 aircraft deployed 21 dropsondes in the southwest sector of an IVT system located between a developing cyclone to the northwest and a major subtropical high pressure system off the California coast (Fig. 6). The IVT makes landfall over northwestern Washington state and Vancouver Island during the first 48 h of the forecast, while subsequent forecast days cover a second heavy precipitation event over the period 1-2 February and the advent of a second IVT in the Gulf of Alaska that is primarily over water during the remainder of the forecast (1-5 February). Improved verifications of WSPD850 and SPCH (Fig. D1a, b), as well IVT (Fig. 33a), generally increase with forecast hour, with the 0-48 h period being relevant to the landfalling system and the 72-168 h period relevant to a new IVT entering the verification domain and making landfall at 168 h.

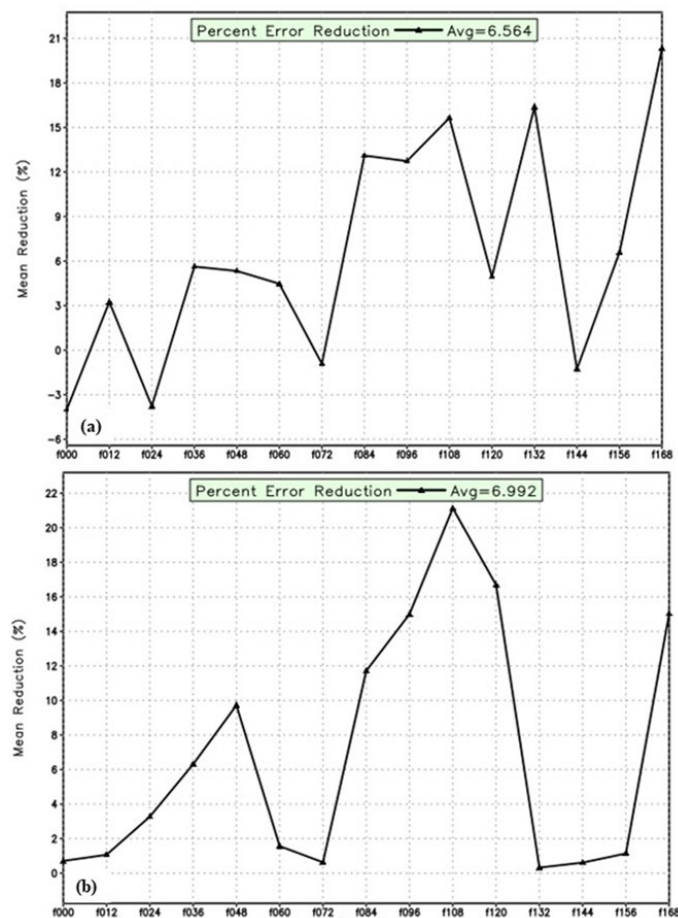


Fig. D1. As in Fig. 33, except for 850 hPa wind speed (a) and 850 hPa specific humidity (b).

The CTRL 12 h forecast IVT MAE reduction (not shown) occurs downstream of the dropsondes deployed 12 h earlier. From 24-48 h, domain-averaged IVT MAE is reduced about 3% due to smaller error growth in the southwest part of the IVT, away from land. However, error growth along the leading edge of the IVT system at the time of maximum precipitation (12 UTC 1 February) is comparable in both CTRL and DENY 84 h forecasts, but the new IVT entering from the west is better captured offshore in CTRL than DENY (Fig. D2a, b), resulting in a net CTRL improvement domain-wide.

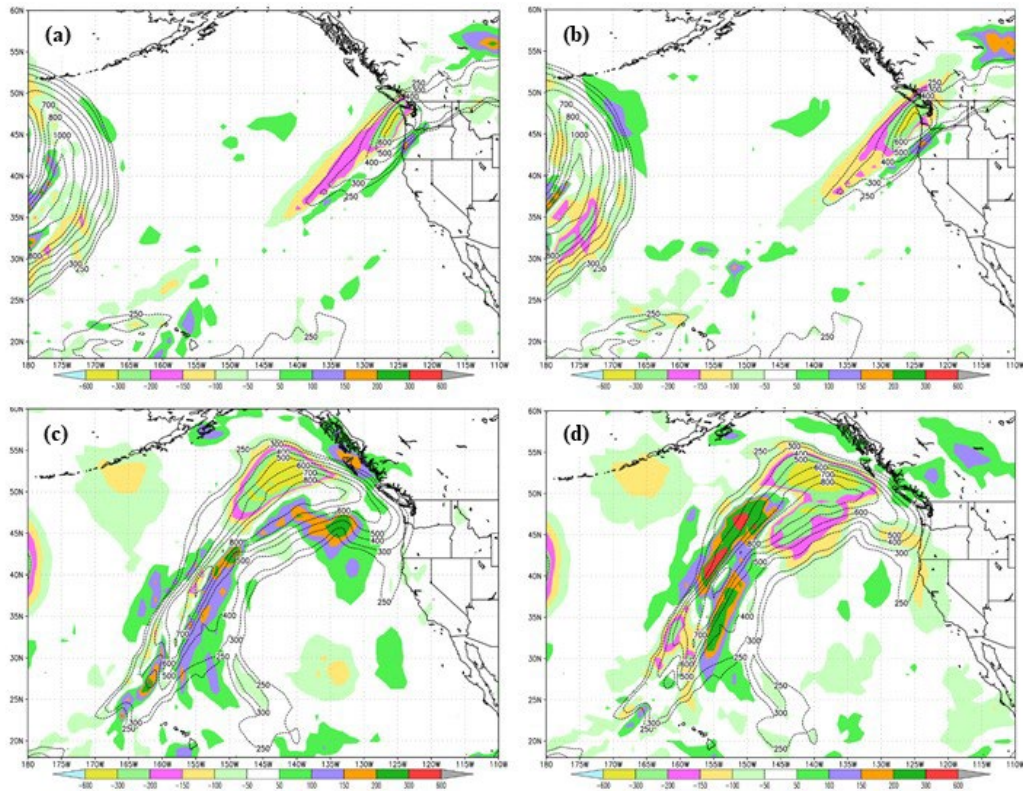


Fig. D2. IOP-2 (29 January) CTRL (a, c) and DENY (b, d) IVT 84 h (a, b) and 168 h (c, d) forecast error (shaded) and ECMO IVT verification analysis contours (black) verifying at 12 UTC 1 February 2020 and 00 UTC 5 February respectively.

As the new IVT propagates east over 72-168 h, the CTRL generally continues to improve over the DENY (Fig. 33a) but, at 00 UTC 5 February when precipitation from the landfalling IVT is beginning to impact Washington state (Fig. 11), the DENY forecast is too weak in the northeast (Fig. D2d) but the CTRL is too strong to the south (Fig. D2c). At this time, the IVT is rotating about a large high-pressure system west of California and an offshore trough with a closed low in the Gulf of Alaska stretches southeast to Oregon (Fig. D3a). The CTRL forecast over develops the trough off the Oregon coast but fails to forecast the closed low in the Gulf of Alaska; the DENY forecast trough is strongly under developed through the entire trough area (Fig. D3b). Within this large-scale pattern, however, both moisture and wind speed errors contribute to the CTRL and DENY IVT forecast errors spatially and also for IVT magnitude. These relationships are discussed below for 850 hPa fields; relationships are similar at both 700 and 925 hPa.

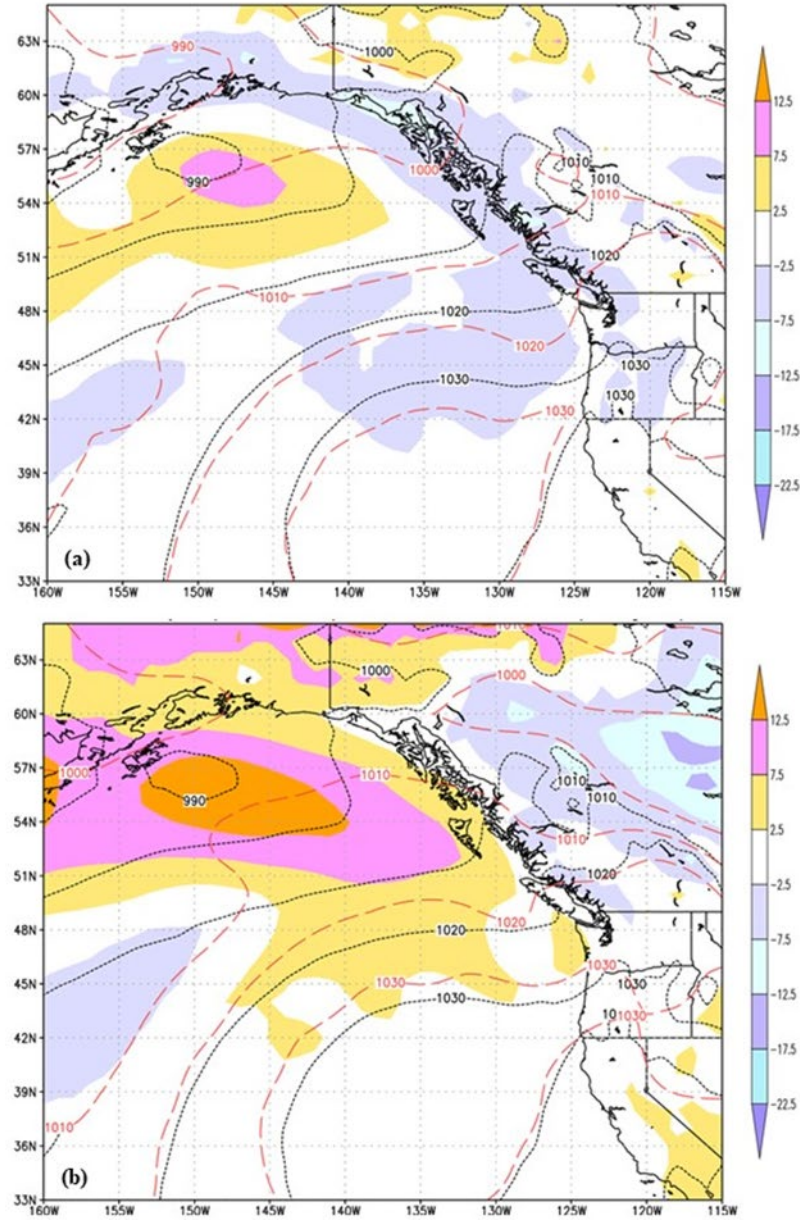


Fig. D3. CTRL (a) and DENY (b) 168 h mean sea-level pressure forecast error (shaded), forecast (red) and verifying analysis for 00 UTC 5 February (black).

The 168 h moisture forecast and error for the CTRL experiment (Fig. D4a) has excess moisture centered at 42 N along 133W, which coincides with an area where the CTRL IVT is too strong, while to the north, centered at 53 N, 140 W, the forecast is too dry and the IVT too weak. Similarly, the DENY experiment is too dry along 53-54 N and the IVT is also too weak. Furthermore, a large area of excess moisture, centered at 46 N 155 W, occurs in the DENY experiment (Fig. D4b) and coincides with the central and southwest IVT values being too large (Fig. D2d); this area has greatly reduced area in the CTRL experiment and is largely responsible for the CTRL IVT error reduction. East-west oriented lower tropospheric wind speed errors along 51 N occupy a much larger area in the DENY experiment than the CTRL (Figs. D5a, b) and also contribute to the large negative IVT forecast errors in this region for both CTRL and DENY (Figs. D2c, d).

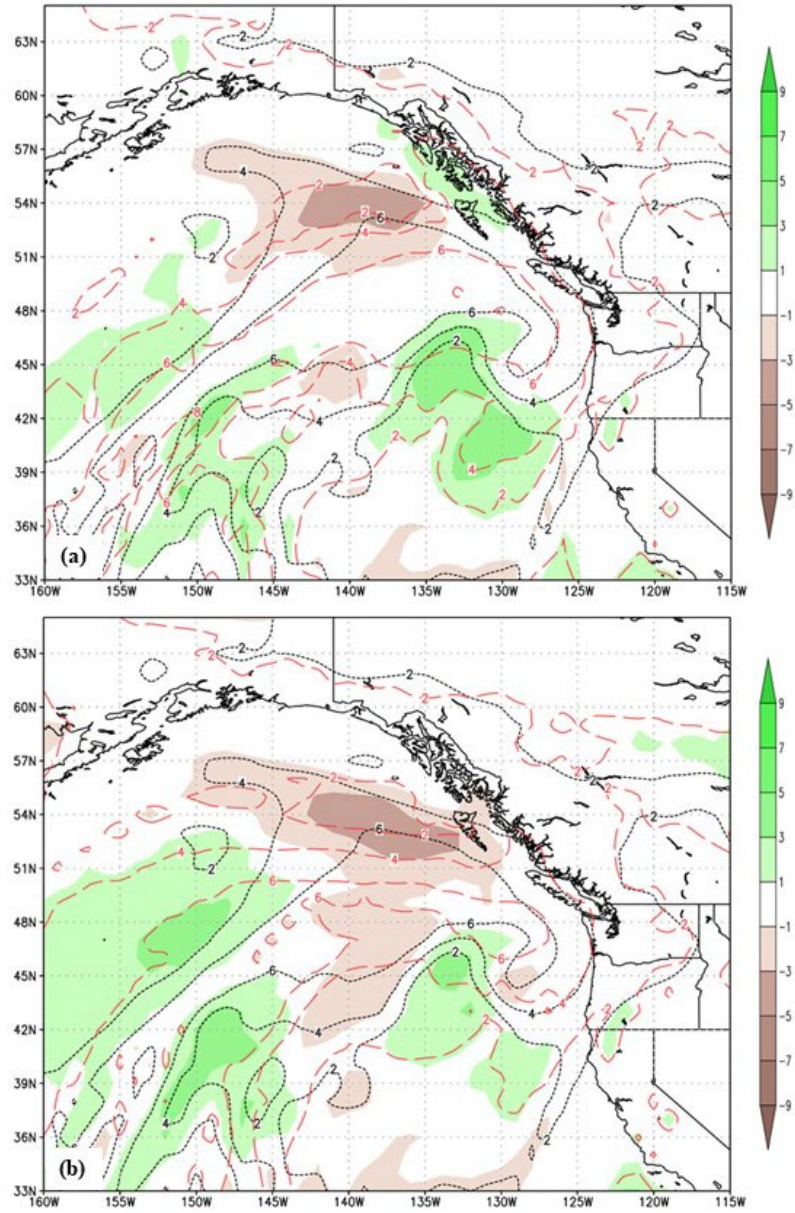


Fig. D4. CTRL (a,) and DENY (b) 168 h specific humidity forecast error at 850 hPa (shaded), forecast contours (red) and verifying analysis for 00 UTC 5 February (black).

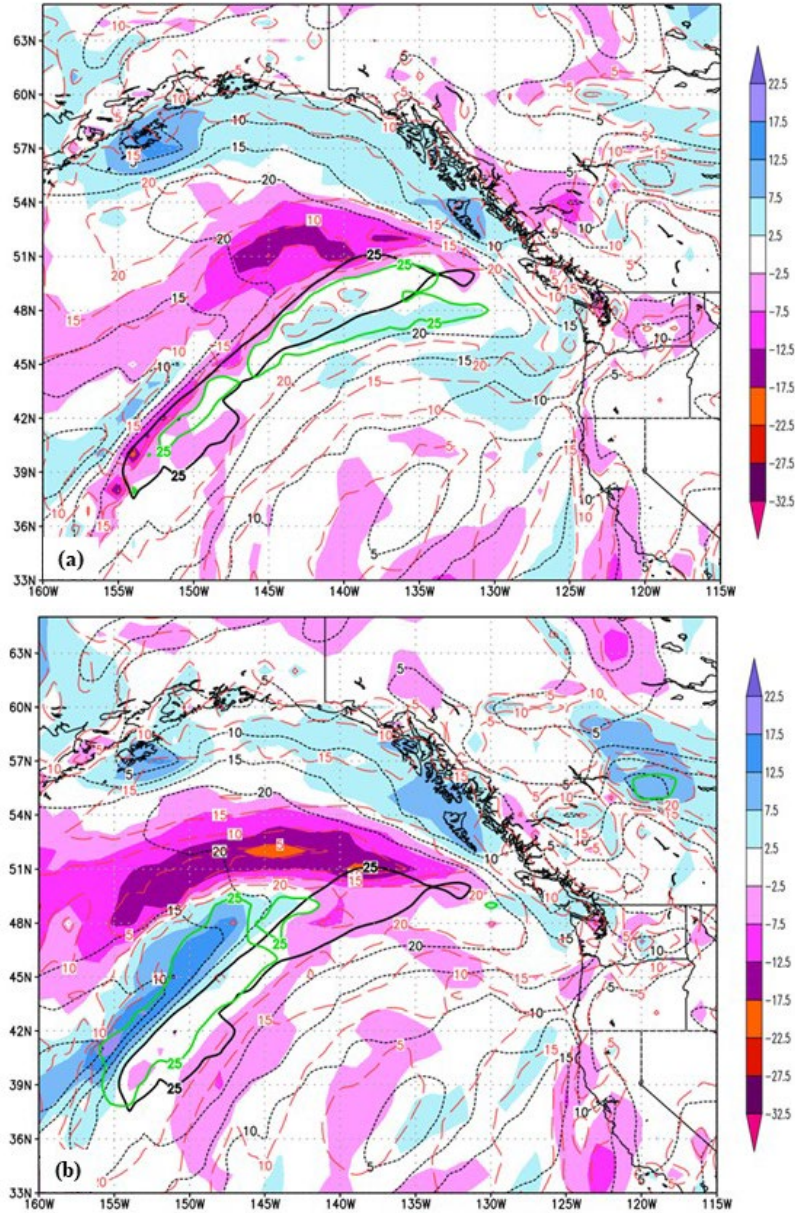


Fig. D5. As in Fig. D4, except for the CTRL (a) and DENY (b) 850 hPa wind speed error (shaded), forecast isotachs (red dash) and verifying analysis isotachs (black dash). The forecast and verification 25 m s^{-1} isotachs are emphasized (green and black thick solid contours respectively).

2. IOPs 8-9

IOP-8 (15 February) is the second of three consecutive IOP days to observe a landfalling IVT offshore from the Pacific Northwest and northern California. Through 60 h, and consistent with the forecast error series for Z700 MAE (Fig. 42) and PMSL MAE (not shown), the CTRL and DENY forecasts for the landfalling IVT are very similar in position and magnitude. Errors for the IVT entering the verification domain from the southwest as the forecast begins are also very similar (not shown). At 72-108 h, the CTRL forecast is improved over the DENY but at 120 h, as a closed low pressure system is generated over Prince William Sound in south Alaska, the associated CTRL errors begin to rise rapidly as the low pressure system extends further southeast along the Canadian coast out to 168 h (Fig. D6a, b).

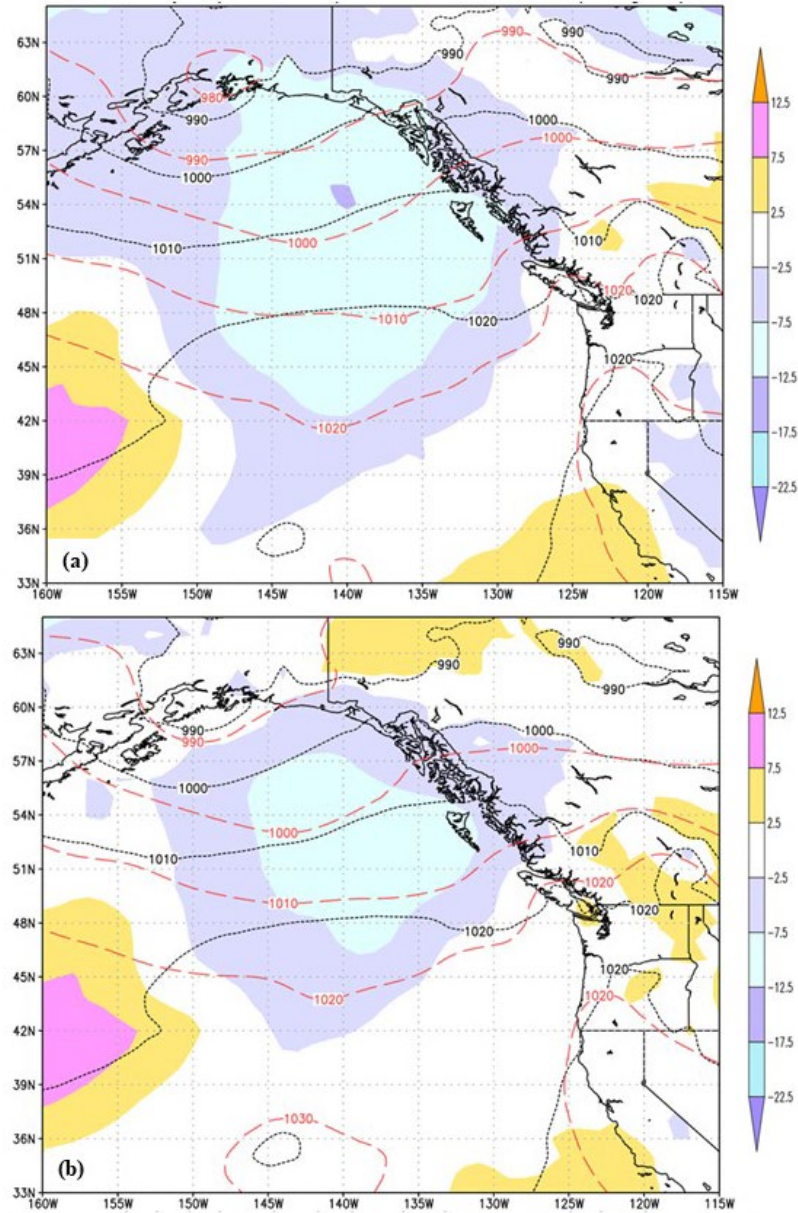


Fig. D6. As in Fig. D3, except for IOP-8 and CTRL (a) and DENY (b) 168 h forecasts with a verification date of 00 UTC 22 February.

For IOP-9 (16 February), initialized 24 h later than the IOP-8 forecast but also verifying on 22 February, the CTRL 144 h forecast (Fig. D7a) is improved over the IOP-8 168 h forecast (Fig. D6a) and the same is true for the DENY pair. However, the error increase from the 144 h to the 168 h forecast is larger for the CTRL (Fig. 7b, d), implying that the CTRL initial conditions contain a potential for a fast-growing error compared to the DENY forecast pair (Table D1).

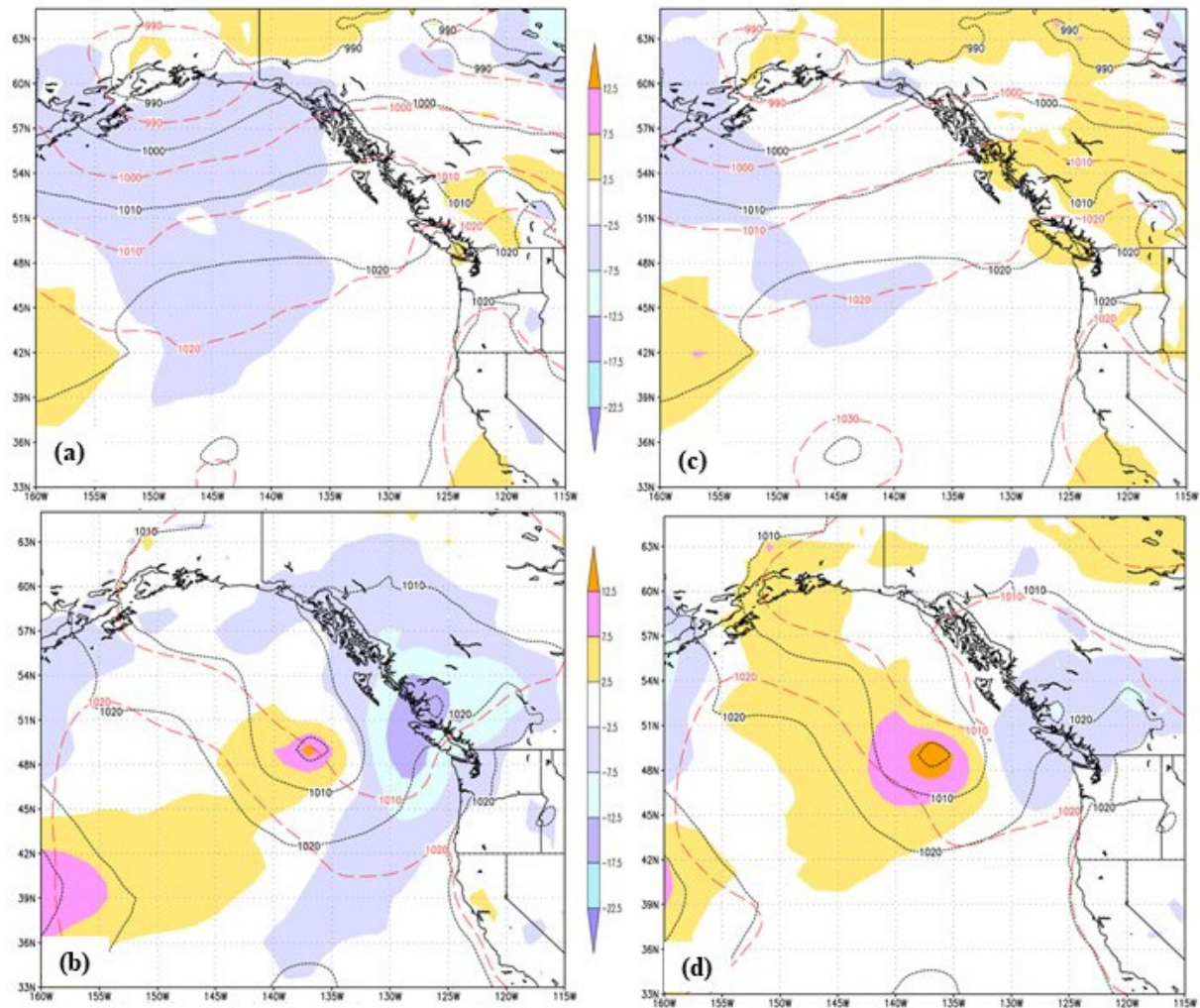


Fig. D7. As in Fig. D3, except for IOP-9 CTRL (a, b) and DENY (c, d) 144 h (a, c) and 168 h (b, d) forecasts with a verification dates of 00 UTC 22 and 23 February respectively.

The IOP-9 CTRL 168 h IVT forecast is degraded by 20.6% compared to the DENY experiment, again primarily due to the southeastward (progressive) error in the low pressure system now verifying at 49 N 137 W (Fig. D7b, d and Table D1). The DENY 168 h forecasts verifying on 22-23 February are approximately the same quality (2.8 versus 2.4 hPa MAE), while the CTRL forecasts are 30% worse (4.3 versus 3.3 hPa). These results again support the possibility of a fast-growing error for IOP-8 (or earlier) assimilated observations (but not necessarily AR OC dropsondes).

IOP	Verification Date	Forecast length (h)	CTRL	DENY
8	00 UTC 22 February	168	4.340	2.851
9	00 UTC 22 February	144	2.254	1.884
9	00 UTC 23 February	168	3.310	2.428

Table D1. Comparison of mean sea-level pressure MAE for 144 and 168 h forecasts for IOP-8 and IOP-9.

3. IOP-13

The 168 h IOP-13 (7 March) forecast covers a major AR system that will landfall over western Canada during the 0-84 h range, a second IVT maximum at the southwest end of the first system that propagates into the IVT verification area and makes landfall over the Aleutians from 60-144 h and a third, complex IVT system interacting with an offshore cyclone (propagating from the Gulf Alaska) while making landfall in Baja California and impacting the SCAN domain (Table 1). Each of these systems is discussed below.

For the 00 UTC 7 March data assimilation cycle, the GIV deployed 30 dropsondes to sample north and east of a small IVT maximum ahead of the major AR system entering the verification domain at 45 N, upstream of the dropsonde deployment (Fig. 22a). Except for the analysis and 12-36 h forecasts, which were positively impacted by the dropsondes, the IOP-13 forecast had a major negative impact from 48-168 h (Fig. 33a) due to these added data and the cycled atmospheric state that had evolved due to the dropsondes since 24 January.

a. Landfall over BC, Canada

Both the dropsonde-sampled system and its successor IVT maximum circulated anticyclonically around the same high pressure system centered at 36 N, 167 W (Fig. 22a). As this anticyclone migrates east-southeast over the next 48 h (00 UTC 9 March), the leading feature disappears but the trailing IVT system remains coherent and propagates eastward along 50 N (Fig. 22b-d) until landfall at 84 h (12 UTC 10 March) NE of Vancouver Island, BC. While there are no available land-based rainfall data analogous to the CCPA for Canada, the CMORPH data indicate heavy rainfall along the Canadian coast from 48-54 N during the period 12 UTC 10 March – 12 UTC 11 March (Fig. D8a-c).

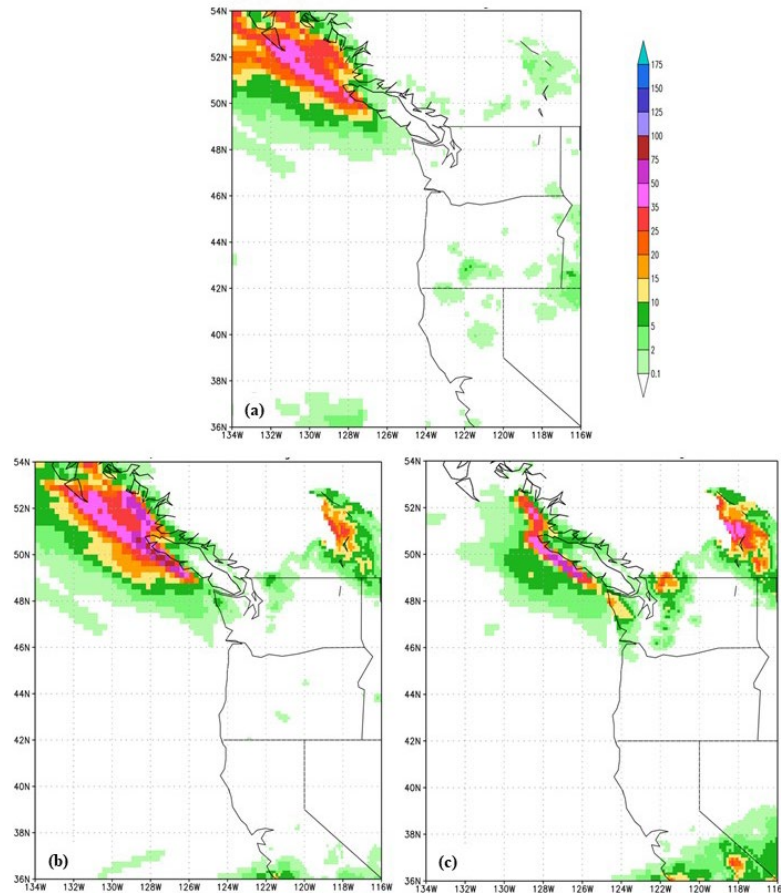


Fig. D8. As in Fig. 5, except for valid date 12 UTC 10 March (a), 00 UTC 11 March (b) and 12 UTC 11 March (c).

Compared to the IVT verifying analysis maximum at landfall (Fig. D9a), the 84 h CTRL IVT system forecast (Fig. D9b) is weaker and spread too far to the north in the northeastern landfalling sector. Offshore, the CTRL forecast system stretches to the west southwest but is, on the whole, south of its verifying position from 140 W to 155 W. The DENY forecast (Fig. D9c) has the same error pattern, but has smaller errors overall and better amplitude and displacement both offshore and at landfall. The negative IVT forecast impact from 60-168 h appears to have contributions from, e.g., 500 hPa geopotential height, 850 hPa wind speed and 850 hPa specific humidity (Figs. D10a-c), all of which are representative of errors at adjacent levels as well.

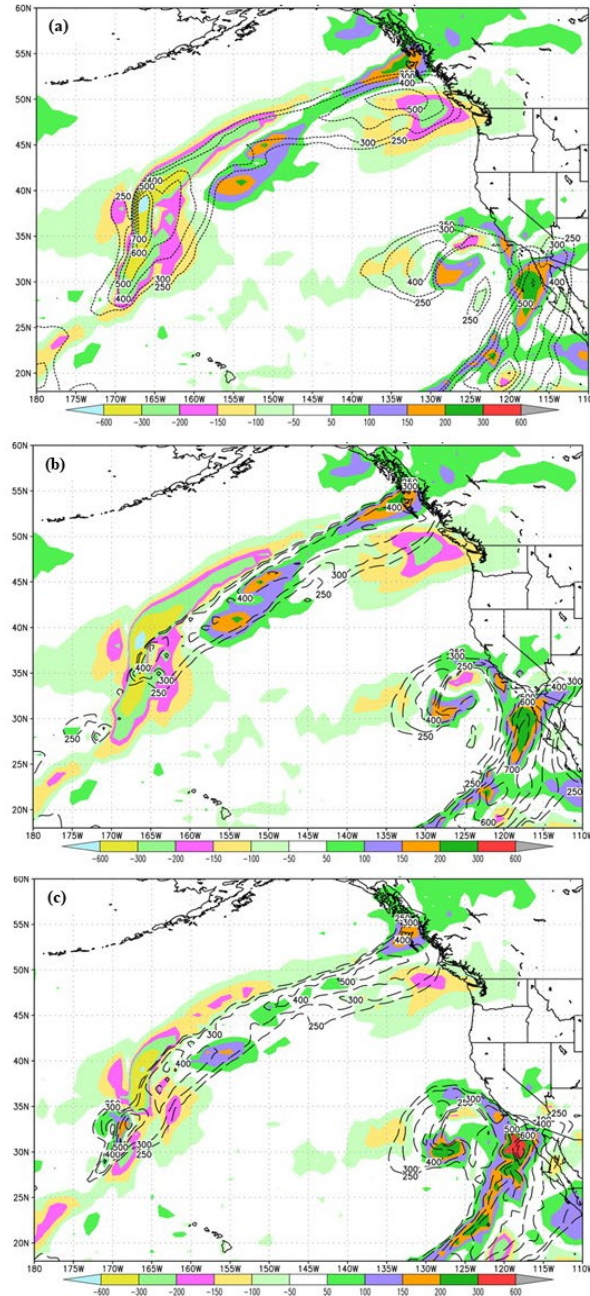


Fig. D9. IOP-13 (7 March) IVT 84 h CTRL forecast error (color) and verifying analysis (black, a), CTRL forecast (black, b) and DENY forecast (black, c) valid on 12 UTC 10 March 2020.

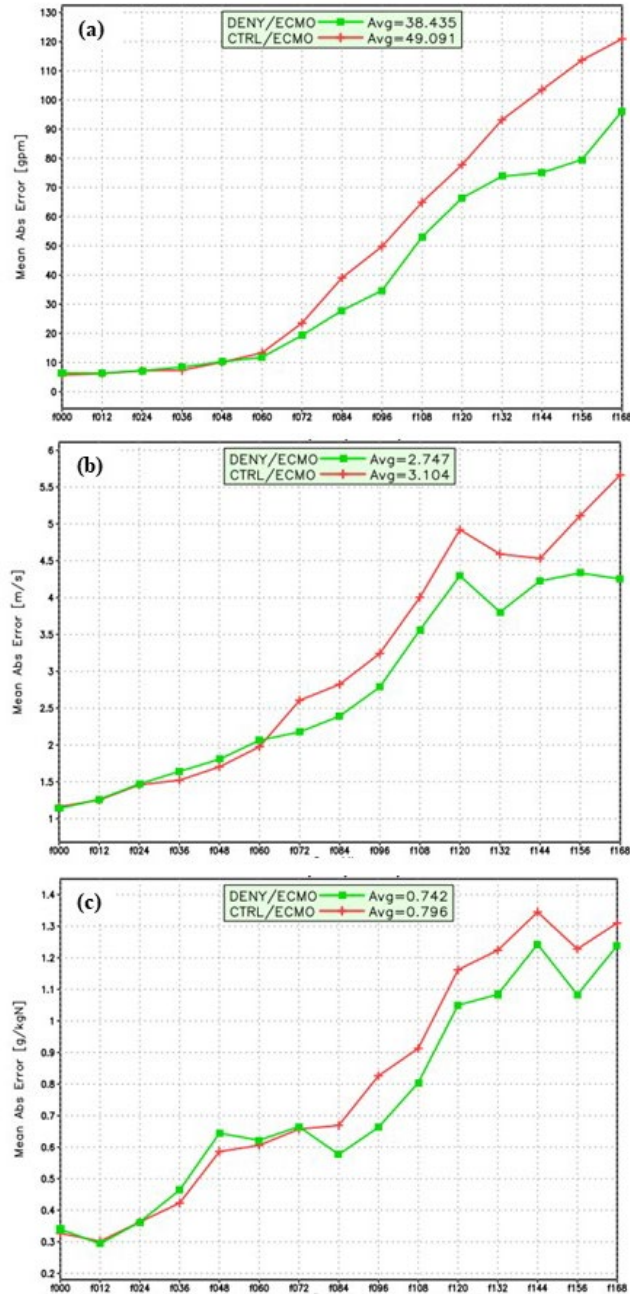


Fig. D10. MAE for 500 hPa geopotential height (a), 850 hPa wind speed (b) and 850 specified humidity (c) for the IOP-13 analysis and 12-168 h CTRL and DENY forecasts initialized at 00 UTC 7 March. The average MAE over the analysis and all forecasts is shown in the legend.

At 84 h, the DENY experiment has positive errors over western Canada and negative errors over the Gulf of Alaska and to the south (Fig. D11a). The CTRL error pattern is similar, but errors are of larger magnitude (Fig. D11b). Similarly, the DENY experiment under predicts the onshore 850 hPa flow south and west of Vancouver Island (Fig. D11c) and the CTRL wind speed errors are larger yet (Figs. D11d). Last, the DENY 850 hPa specific humidity on the north edge of the landfalling IVT at the Canadian coast (53 N) is slightly high (Fig. D11e) and under- predicted on the south side (48 N), consistent with the IVT forecast errors (Figs. D9a-c), and the CTRL forecast has the same pattern but stronger (Fig. D11f).

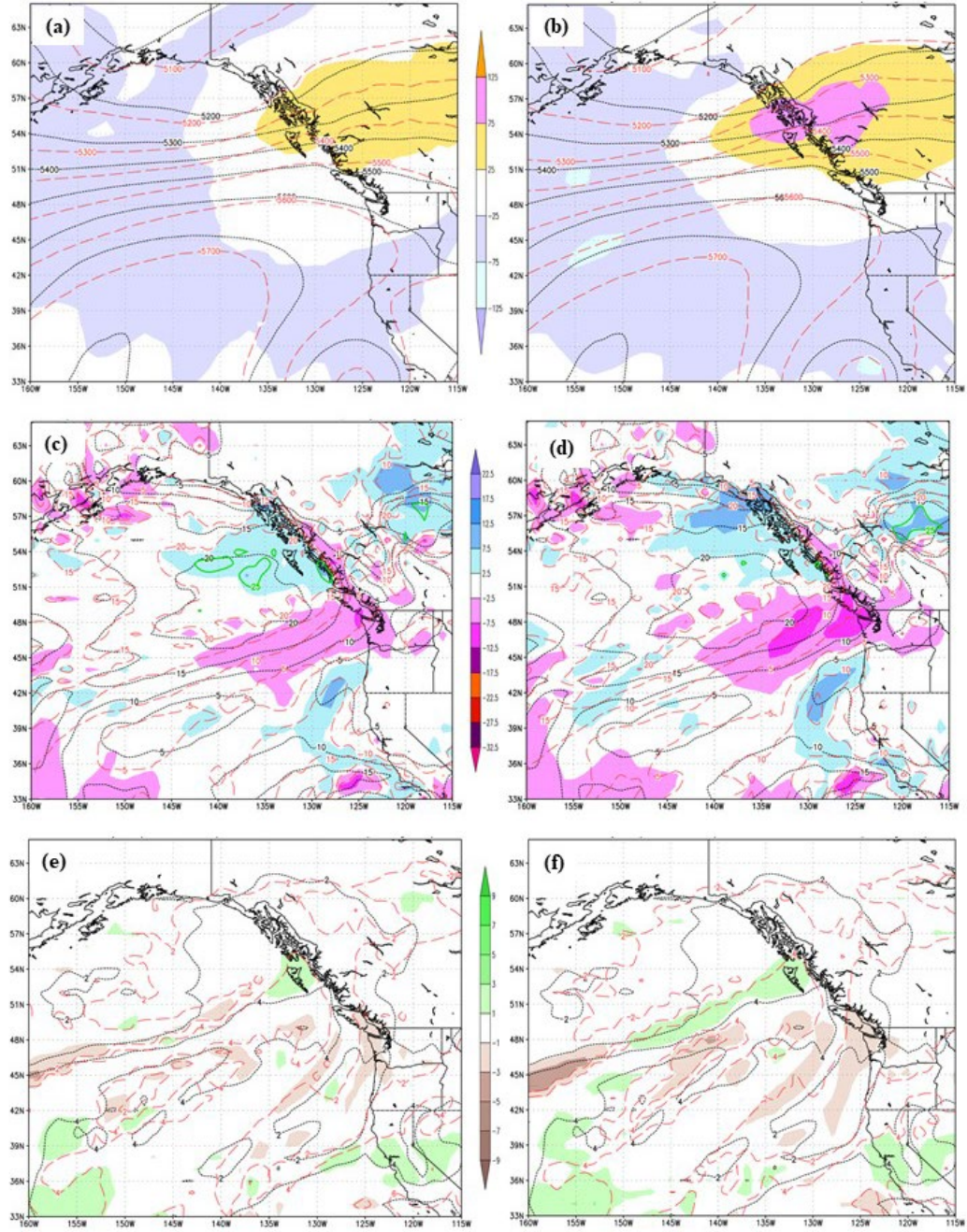


Fig. D11. DENV (a, c, e) and CTRL (b, d, f) 84 h 500 hPa geopotential height (a, b), 850 hPa wind speed (c, d) and 850 hPa specific humidity (e, f) forecast error (shaded), forecast (red) and verifying analysis (black) for IOP-13 initialized at 00 UTC 7 March.

b. Landfall over the Aleutians

The IOP-13 MAE error reductions (Fig. 33b) from 48-168 h reflect larger CTRL errors, relative to DENV, in predicting the complex, lower-level (e.g., 850 hPa) mid-latitude ridge-trough distribution over the verification domain and particularly in the sub-tropical mid-Pacific east of the date line (Fig. D12a). Southerly flow with enhanced moisture from the tropics, rotating about the high pressure center at 36 N 167 W, is well forecast from 12-24 h, consistent with the error statistics (Fig. 33b), but the southerly flow is weakened and rotated to easterly flow at 36 h in the

vicinity of 26-29 N 170-173 E (Fig. D12b). The CTRL IVT maximum at this time is substantially under-predicted as consequence due to the divergent easterly flow and weaker wind speeds; moreover, the accompanying low pressure system, immediately west, is also under-predicted (not shown). The DENY IVT scenario, while still under-predicted at 36 h, lacks the easterly flow and, hence, maintains southerly flow along the western side of the high pressure center, more similar to the verifying analysis (Fig. D12d). Moreover, the accompanying DENY forecast low pressure system continues to be less under-predicted at 48-60 h and beyond.

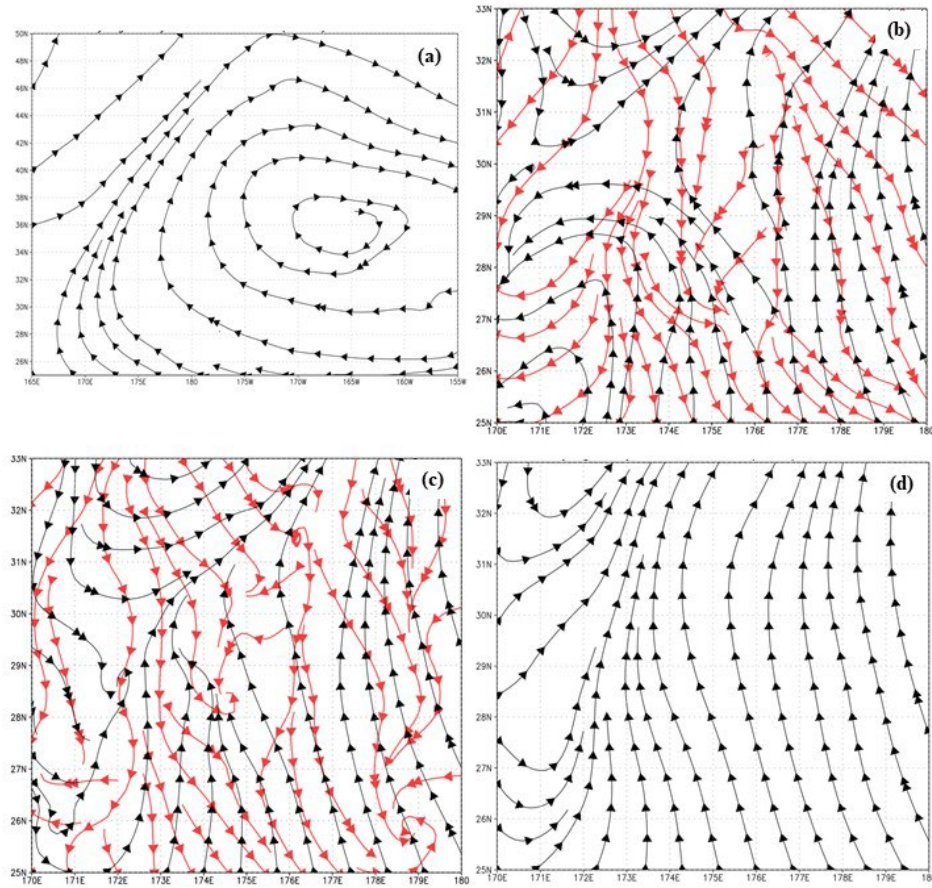


Fig. D12. 850 hPa streamlines for the ECMO verifying analyses at the IOP-13 initial date (00 UTC 7 March) and, in a close-up view, at the verification date for the 36 h forecast (d). CTRL (b) and DENY (c) 36 h forecast (black) and error streamlines (red) indicate a weaker tropical moisture flux at 850 hPa.

From 00 UTC 10 March to 12 UTC 12 March (72-132 h in the IOP-13 forecast), the observed low pressure area associated with the north-eastward propagating IVT is absorbed into a mid-latitude trough, intensifies and continues propagating north, with the IVT gradually wrapping around cyclonically (Figs. 22d-e, D13). However, in both the CTRL and DENY forecasts, beginning at 96 h (00 UTC 11 March), the IVT propagates northeastward, following the continuing trajectory of the AR into Canada as opposed to making landfall in the Aleutians (Figs. 45, 46a, b). A surface low pressure center develops under the propagating IVT at 108 h and the low pressure anomaly/error continues to develop (Figs. D14a-c) as it rotates anticyclonically around the large-scale high pressure region centered at 41 N 140 W. At landfall (132 h, Fig. 14c), the Gulf of Alaska and northeast Pacific is dominated by a massive negative error anomaly and a low pressure

center vice an anticyclone west of Vancouver, BC and the Pacific Northwest (Fig. D13). The DENV forecast follows much the same pattern but the low pressure system is not as strong (Figs. D14d-f).

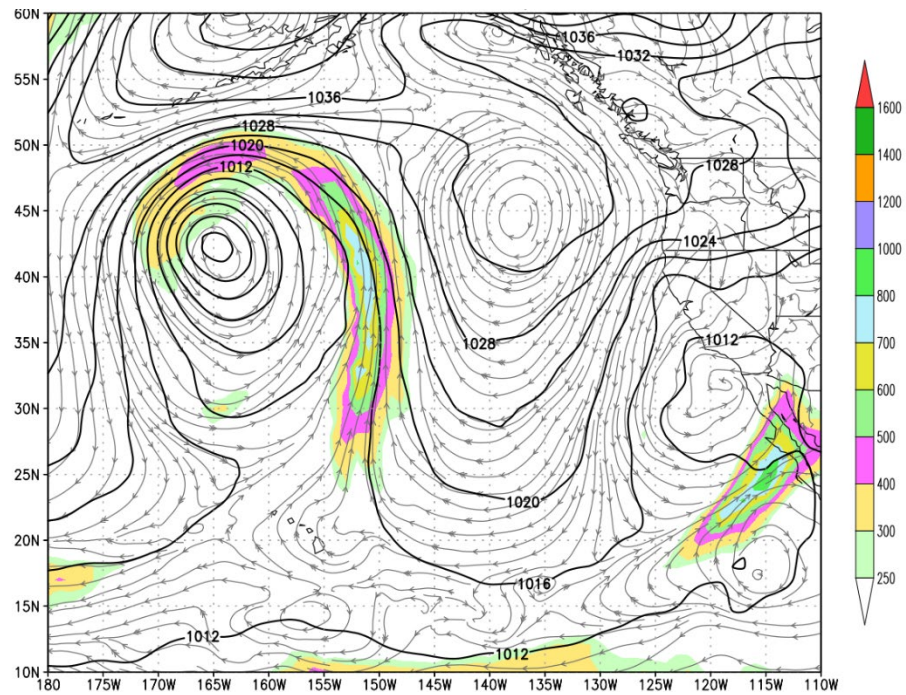


Fig. D13. As in Fig. 3 except for 12 UTC 12 March 2020.

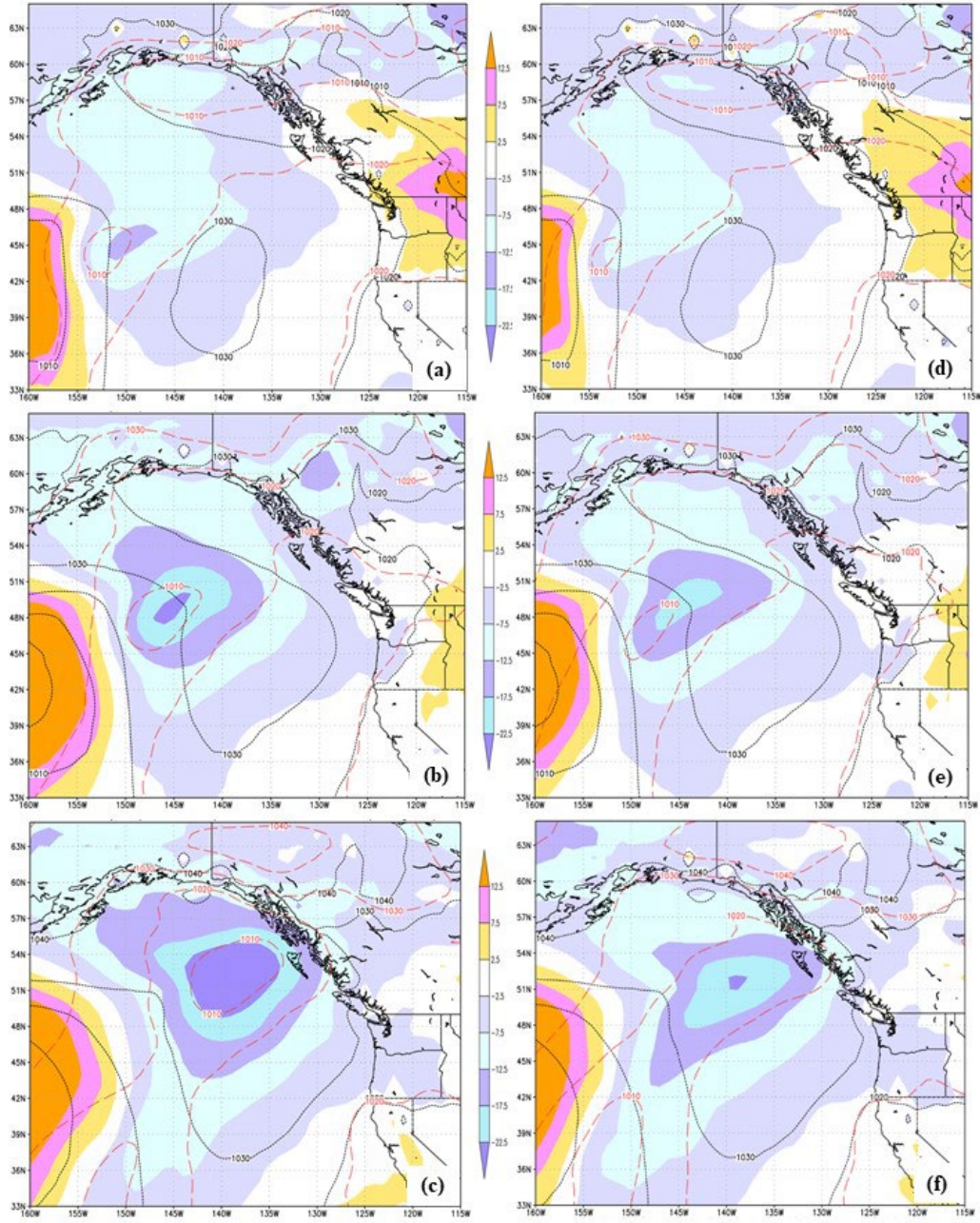


Fig. D14. IOP-13 CTRL (a, b, c) and DENY (d, e, f) 108 (a, d), 120 (b, e) and 132 (c, f) h mean sea-level pressure forecast error (shaded), forecast (red) and verifying analysis (black) at 12 UTC 11 March.

APPENDIX E

Precipitation Statistics for all AR2020 OC Cases and All IOPs

1. Full OC statistical calculations

The I_{DFpstd} statistic measures the variability in difference between the forecast and CCPA verification, with the former being interpolated to the verifying observed grid. In cases of minimal observed 24 h accumulated precipitation (domain-average CCPA < 0.002 mm/day), I_{DFpstd} (and other precipitation statistics) are not calculated. Furthermore, in cases when CTRL-DENY differences and the CCPA observation produce unrealistic improvement values (e. g., $|I_{DFpstd}| > 100\%$), these statistics are also not calculated.

2. I_{DFpstd} statistic for the PNNC domain

Precipitation impacts for a specific statistic are summarized by plotting precipitation impact categories (Fig. 14a) for each of the 96 cases and each forecast hour (Fig. E1). For I_{DFpstd} , impacts vary across consecutive cases and by forecast hour for a particular case. The maximum positive impact across all forecast hours is 15.9% for the 12 UTC 19 February initialization and the maximum negative impact is -20.6% for 00 UTC 6 March. In neither of these cases is the initialization on an IOP date, so that most of the overall impact is governed by the evolution of the wind and moisture environment through the cycled data assimilation system. A time series of the forecast-averaged I_{DFpstd} impact (Fig. E1b) shows periods of mostly positive average impact that include IOPs (e.g., 00 UTC 10 February to 12 UTC 19 February, IOPs 7-9) and also periods of mostly negative impact, also including IOPs (e.g., 12 UTC 27 January to 00 UTC 5 February, IOPs 2-5).

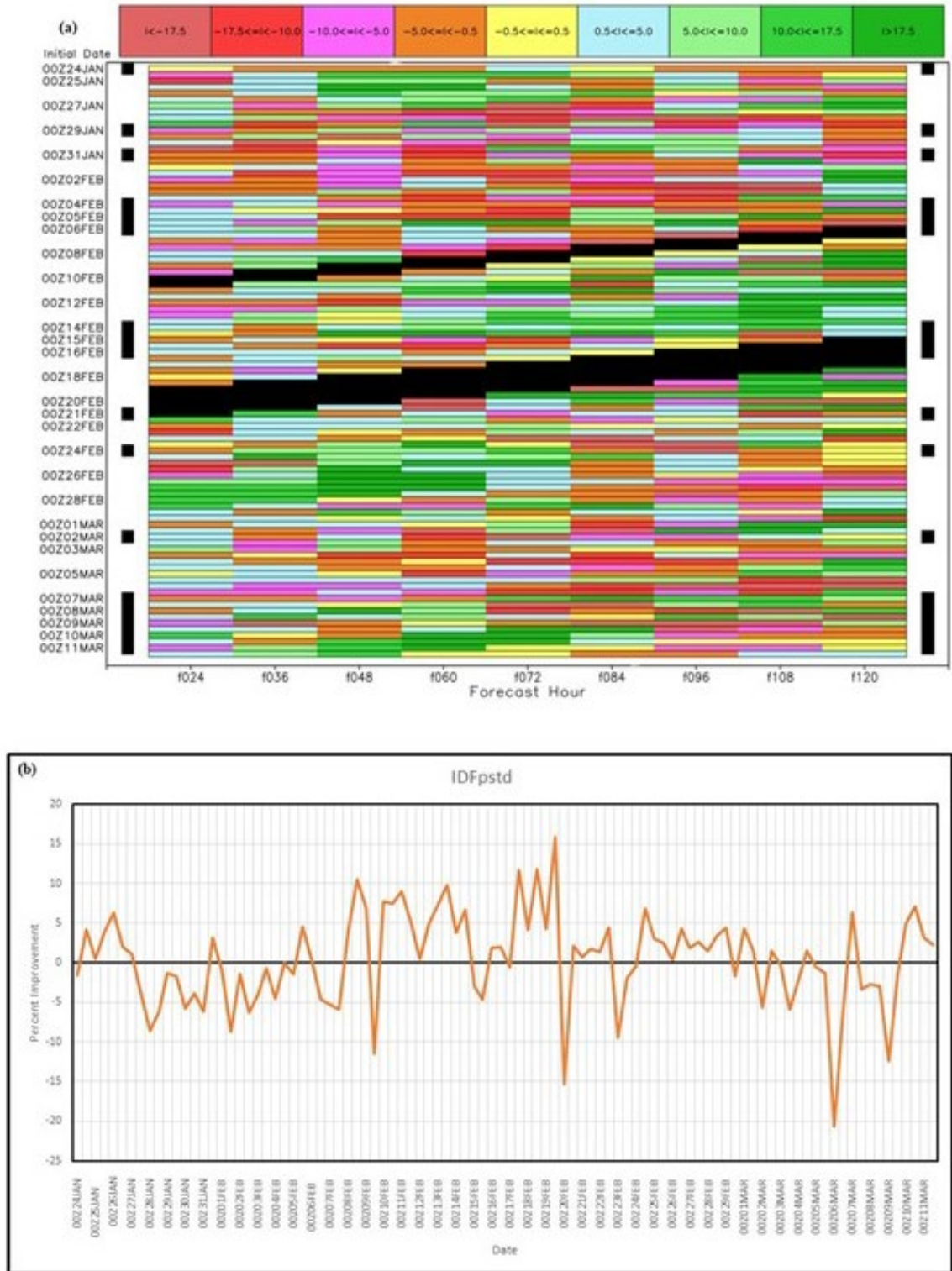


Fig. E1. Percent improvement (a) for the ID_{fpstd} statistic in the PNNC domain (Table 1) for forecast hours 24-120 and initialization times every 12 h from 00 UTC 24 January to 12 UTC 11 March. Black areas are discarded cases due to minimal observed 24 h accumulated rainfall in the domain (Fig. 2b). Forecast-averaged improvements (b) are shown as a time series over the same period.

3. SCAN domain statistics

SCAN domain statistics (Fig. E2a) generally show impacts of <1%, except for isolated 3% positive impact for 48, 60 and 96 h and 5-8% degradation at 60 and 120 h for domain maximum values. At 108-120 h, however, there are unusually large positive impacts to the maximum precipitation difference ($ID_{F_{pmax}}$), accompanied by a large positive percentage of improved cases (Fig. E2b), which indicates that the CTRL has a better match of heavy precipitation locations to the observed than the DENY. Overall, however, the CTRL has fewer positive impacts for most statistics than the DENY.

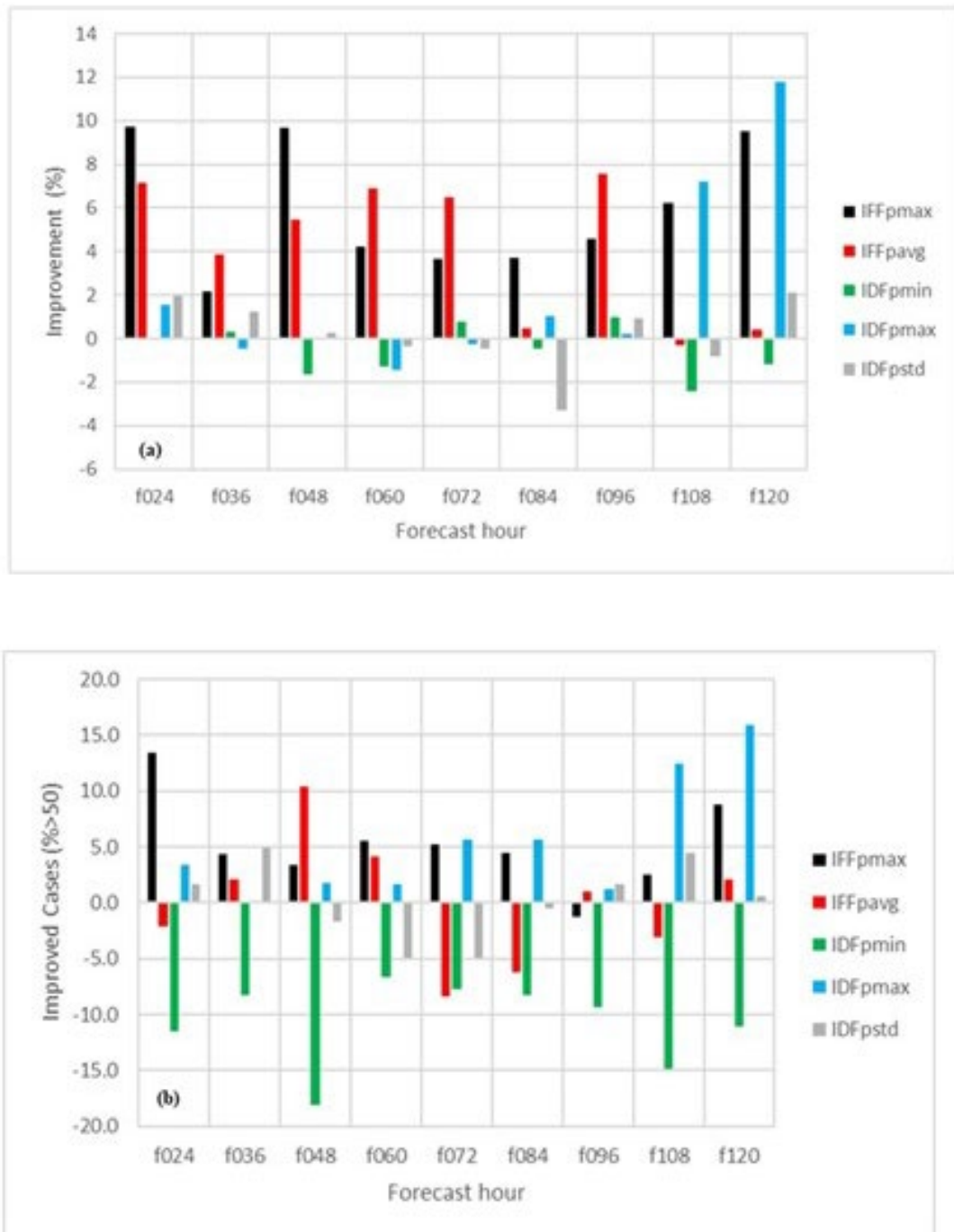


Fig. E2. As in Fig. 53, except for the SCAN domain.

4. IOP only statistics

Precipitation impact statistics for all forecast hours and each IOP are summarized in Fig. E3. Very few IOPs have positive impact for more than two statistical measures. For 24 h forecasts, improvements and degradations are scattered across different IOPs with little sustained impact.

Coherence across IOPs increases from 36-60 h, the latter forecasts having by far the strongest and most coherent improvements across the statistical measures for both PNNC and SCAN domains. On the other hand, for 72-84 h, improvements characterize the PNNC domain while degradations dominate for the SCAN domain. From 96-132 h, the DFpstd statistic is predominantly improved (see also Fig. 55) as much as 10-20%, indicating a better overall fit of forecast precipitation with the verification over the SCAN domain.

Forecast Hour	IOP	Domain	Valid Dates	FFpmax	FFpavg	DFpmin	DFpmax	DFpstd
24	1	pnnc	4	1.09	2.67	-4.05	9.85	0.41
	2	pnnc	4	-8.35	0.12	1.75	-23.09	-4.67
	3	pnnc	2	10.42	3.97	-3.75	8.69	3.98
	5	pnnc	3	0.46	1	2.06	-4.22	0.63
	6	pnnc	2	-0.72	0.77	2.81	-6.26	0.59
	8	pnnc	1	7.87	1.97	22.64	18.76	0.75
	12	scan	2	6.36	-1.97	0.63	-18.94	0.56
	16	scan	4	-6.66	1.15	-2.8	17.96	-0.36
	17	scan	3	-7.75	0.84	-3.31	15.65	-0.53
36	1	pnnc	4	3.96	-0.23	2.42	-9.18	0.56
	2	pnnc	3	-7.25	0.59	2.71	-20.89	-5.38
	3	pnnc	1	-3.87	-3.01	2.49	8.77	-1.74
	4	pnnc	4	4.95	2.75	2.07	-2.17	-3.97
	5	pnnc	3	6.38	3.16	-1.86	-9.82	-5.38
	6	pnnc	2	2.98	0.85	0.52	3.17	-0.93
	8	pnnc	1	3.29	-2.8	1.04	-1.89	-2.09
	10	scan	1	13.78	12.97	-0.43	-6.32	3.35
	12	scan	2	-5.08	-1.23	0.98	27.87	1.64
	15	scan	5	-4.79	-2.06	0.19	25.75	0.74
	16	scan	4	11.61	-3.35	0.34	32.67	2.19
	17	scan	3	-12.71	-4.68	0.64	34.58	2.68
48	1	pnnc	3	-5.03	-0.15	2.37	15.9	10.03
	2	pnnc	3	-11.56	-4.34	8.61	-22.38	-10.74
	4	pnnc	3	-3.38	-2.43	0.89	4.8	3.04
	5	pnnc	2	-0.35	-4.68	0.04	1.72	4.33
	6	pnnc	1	2.49	1.92	2.36	6.65	-0.95
	7	pnnc	1	-3.75	3.62	34.69	34.66	19.04
	10	scan	1	6.41	11.46	-0.36	-18.98	3.76
	12	scan	2	12.73	-3.55	0.9	18.29	4
	15	scan	4	-0.74	-6.06	-2.83	-19.11	0.6
	16	scan	3	3.04	-4.46	-3.61	-17.81	-0.52
	17	scan	2	5.06	-3.88	-1.8	-28.98	-0.6
60	1	pnnc	3	-8.02	1.88	1.09	26.67	12.24
	2	pnnc	2	14.04	0.51	-2.99	1.06	-1.57
	4	pnnc	3	7.33	-12.79	-3.97	5.87	5.94
	5	pnnc	2	4.43	4.25	-6.54	23.03	14.38
	6	pnnc	1	3.88	5	14.34	11.96	4.06
	7	pnnc	1	14.61	-0.03	12.55	39.34	22.2
	12	scan	1	39.33	5.65	3.07	-81.47	2.07
	14	scan	5	-12.27	-0.74	1.29	7.25	1.32
	15	scan	4	-20.1	-3.19	2.22	4.61	1.72
	16	scan	3	-5.87	4.53	2.51	10.53	1.43
	17	scan	2	-5.34	4.53	1.72	4.27	1.36
72	1	pnnc	2	-9.88	0.68	1	7.22	10.94
	2	pnnc	2	-8.12	3.59	-9.08	-3.03	-0.08
	4	pnnc	2	0.36	10.31	19.16	-0.35	2.67
	5	pnnc	1	34.66	21.05	82.81	35.46	29.45
	12	scan	1	6.52	13.18	-1.23	0.68	1.58
	14	scan	4	-5.29	5.72	0.21	-7.68	-2.2
	15	scan	3	-8.61	0.52	-0.14	-6.19	-3.27
	16	scan	2	7.76	3.69	-0.2	-11.45	-5.51
	17	scan	1	7.11	8.19	-1.35	-12.07	-11.04
84	1	pnnc	2	3.19	0.55	-3.7	-4.43	-0.33
	2	pnnc	1	4.65	9.97	-1.36	0.34	1.11
	4	pnnc	2	26.9	15.1	-1.13	-38.41	9.05
	5	pnnc	1	1.64	-2.86	25.27	-6.53	7.54
	13	scan	5	-0.48	3.61	-0.08	10.79	-1.07
	14	scan	4	-3.68	2.81	-0.14	12.3	-3.41
	15	scan	3	-6.26	-4.07	-2.04	7.62	-1.96
	16	scan	2	-8.66	-7.19	-2.91	1.73	-2.18
	17	scan	1	1.15	12.62	-12.8	7.42	-3.68
96	1	pnnc	1	14.18	-2.47	0.59	15.13	4.89
	2	pnnc	1	6.64	16.71	24.26	22.99	8.55
	4	pnnc	1	15.59	26.74	1.77	-77.48	-18.11
	13	scan	4	0.21	1.47	2.91	-11.5	3.64
	14	scan	3	2.89	-0.3	3.19	11.4	3.49
	15	scan	2	4.28	-2.82	-0.45	-15.74	1.77
	16	scan	1	10.44	-13.84	-2.75	21.1	2.3
108	1	pnnc	1	-3.88	-4.15	1.94	7.91	-3.53
	4	pnnc	1	7.16	-6.14	18.1	26.42	-4.34
	13	scan	4	-6.78	-0.52	0.09	5.2	-0.08
	14	scan	3	-2.62	-6.77	-0.8	10.01	2.15
	15	scan	2	-2.83	3.79	0.46	11.25	3.73
	16	scan	1	1.73	11.13	-0.03	19.44	7.24
120	13	scan	3	7.53	3.7	-1.31	7.85	7.73
	14	scan	2	-0.27	11.76	-0.52	-8.2	7.09
	15	scan	1	-4.5	7.03	0.51	-0.1	20.79
132	13	scan	3	-2.78	1.97	-0.76	0.71	0.83
	14	scan	2	-5.15	-6.45	3.16	3.88	3.63
	15	scan	1	-3.5	2.75	11.6	-3.17	10.03

Fig. E3. Summary of precipitation improvement statistics (%) for 24-36 (a), 48-60 (b), 72-84 (c), and 96-120 (d) h forecasts, including all contributing IOPs in the AR2020 OC. The verification domain and number of valid dates for each IOP are tabulated. Color key is as in Fig. 54. Missing IOPs do not have forecasts that verify at times of impactful precipitation.

REFERENCES

- Aberson, S., Cione, J., Wu, C.-C., Bell, M., Halverson, J., Fogarty, C., Weissmann, M. 2010. Aircraft Observations of Tropical Cyclones. Global Perspectives on Tropical Cyclones: From Science to Mitigation (pp.227-240) Chan J.C.L. and Kepert J. (Eds). DOI:10.1142/9789814293488_0008.
- Ancell, B. C., A. Bogusz, M. J. Lauridsen, and C. J. Nauert, 2018: Seeding chaos: The dire consequences of numerical noise in NWP perturbation experiments. *Bull. Amer. Meteor. Soc.*, **99**, 615–628, <https://doi.org/10.1175/BAMS-D-17-0129.1>.
- Brennan, M.J., Kleist, D., Howard, K., and Majumdar, S.J., 2015: The Impact of Supplemental Dropsonde Data on the Structure and Intensity of Tropical Storm Karen (2013) in the NCEP Global Forecast System. *Wea. Forecasting*, **30**, 683-691.
- Elless, T.J., Wu, X., and Tallapragada, V., 2021. Identifying Atmospheric River Reconnaissance Targets Using Ensemble Forecasts. The 2021 Blue Book. 1-07. http://bluebook.meteoinfo.ru/uploads/2021/sections/BB_21_S1.pdf.
- Harris, L.M., and Lin, S.-J., 2013. A two-way nested global-regional dynamical core on the cubed-sphere grid. *Mon. Wea. Rev.*, **141**(1), 283– 306.
- Harris, L., Zhou, L., Chen, X., and Chen, J.-H., 2020a. The GFDL finite-volume cubed-sphere dynamical core. NOAA Technical Memorandum. <https://doi.org/10.25923/7h88-c534>
- Harris, L., Zhou, L., Lin, S.-J., Chen, J.-H., Chen, et al., 2020b. GFDL SHiELD: A unified system for weather-to-seasonal prediction. *Journal of Advances in Modeling Earth Systems*, **12**, e2020MS002223. <https://doi.org/10.1029/2020MS002223>.
- Hodyss, D., and S. J. Majumdar, 2007: The contamination of ‘data impact’ in global models by rapidly growing mesoscale instabilities. *Quart. J. Roy. Meteor. Soc.*, **133**, 1865–1875, <https://doi.org/10.1002/qj.157>.
- Hou, D., Charles, M., Luo, Y., Toth, Z., Zhu, Y., Krzysztofowicz, R., Lin, Y., Xie, P., Seo, D., Pena, M., and Cui, B., 2014. Climatology-calibrated precipitation analysis at fine scales: Statistical adjustment of stage IV toward CPC gauge-based analysis. *J. Hydrometeor.*, **15**, 2542–2557, doi: 10.1175/JHM-D-11-0140.1.
- Jolliffe, I. T., and D. B. Stephenson, 2003: *Forecast Verification: A Practitioner’s Guide in Atmospheric Science*. John Wiley and Sons, 240 pp.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., and Xie, P., 2004. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeor.*, **5**, 487–503, doi: [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009: Improving incremental balance in the GSI 3DVAR analysis system. *Mon. Wea. Rev.*, **137**, 1046–1060, <https://doi.org/10.1175/2008MWR2623.1>.
- Kleist, D. T., and Ide, K., 2015a. An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-Hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, doi:10.1175/MWR-D-13-00351.1.

- Kleist, D. T., and Ide, K., 2015b. An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Mon. Wea. Rev.*, 143, 452–470, doi:10.1175/MWR-D-13-00350.1.
- Kleist, D.T., R. Treadon, C. Thomas, H. Liu, K. Bathmann, D. Merkova, C. R. Martin, H. Shao, L. Cucurull, J. G. Yoe, and V. S. Tallapragada, 2021:. NCEP Operational Global Data Assimilation Upgrades: From Versions 15 through 16. In 101st American Meteorological Society Annual Meeting. AMS.
- Lavers, D. A., Rodwell M.J., Richardson D.S., Ralph F.M., Doyle J.D., Reynolds C.A., Tallapragada V., and Pappenberger F., 2018. The gauging and modeling of rivers in the sky. *Geophys. Res. Lett.*, 45, 7828–7834, <https://doi.org/10.1029/2018GL079019>.
- Lavers, D.A., Ingleby, N.B., Subramanian, A.C., Richardson, D.S., Ralph, F.M., Doyle, J.D., Reynolds, C.A., Torn, R.D., Rodwell, M.J., Tallapragada, V. and Pappenberger, F., 2020. Forecast errors and uncertainties in Atmospheric Rivers. *Wea. Forecasting*, 35(4), 1447–1458.
- Lin, S.-J., 2004. A “vertically Lagrangian” finite-volume dynamical core for global models. *Mon. Wea. Rev.*, 132(10), 2293–2307.
- Lin, S.-J., and Rood, R.B., 1997. An explicit flux-form semi-Lagrangian shallow-water model on the sphere, *Q. J. R. Meteorol. Soc.*, 123, 2477–2498.
- Lin, Y. and Mitchell, K.E., 2005. The NCEP Stage II/IV hourly precipitation analyses: development and applications. Preprints, 19th Conf. on Hydrology, American Meteorological Society, San Diego, CA, 9-13 January 2005, Paper 1.2.
- Lord, S., Wu, X., Tallapragada, V. and Ralph, F. M., 2022a. The Impact of Dropsonde Data on the Performance of the NCEP Global Forecast System During the 2020 Atmospheric Rivers Observing Campaign. Part 1: Precipitation. Submitted to *Weather and Forecasting*.
- Lord, S., Wu, X. and Tallapragada, V., 2022b. The Impact of Dropsonde Data on the Performance of the NCEP Global Forecast System During the 2020 Atmospheric Rivers Observing Campaign. Part 2: Dynamic variables and humidity. Submitted to *Weather and Forecasting*.
- McCormack, J.P., Eckermann, S.D., Siskind, D.E., and McGee, T.J., 2006. CHEM2D-OPP: A new linearized gas-phase ozone photochemistry parameterization for high-altitude NWP and climate models, *Atmos. Chem. Phys.*, 6, 4943–4972.
- McCormack, J.P., Hoppel, K.W., and Siskind D.E., 2008. Parameterization of middle atmospheric water vapor photochemistry for high-altitude NWP and data assimilation. *Atmos. Chem. Phys.*, 8, 7519–7532.
- Majumdar, S.J., Brennan, M.J., and Howard, K., 2013: The Impact of Dropsonde and Supplemental Rawinsonde Observations on Track Forecasts for Hurricane Irene (2011). *Wea. Forecasting*, 28, 1385–1403.
- Putman, M. and Lin, S.-J., 2007. Finite-volume transport on various cubed-sphere grids. *J. Comp. Phys.*, 227, 55–78.
- Ralph, F. M., et al., 2014: A vision for future observations for western U.S. extreme precipitation and flooding. *J. Contemp. Water Res. Educ.*, 153, 16–32, <https://doi.org/10.1111/J.1936-704X.2014.03176.X>.

- Ralph, F. M., Dettinger, M.D., Cairns, M.M., Galarneau, T.J., and Eylander, J., 2018. Defining “atmospheric river”: How the Glossary of Meteorology helped resolve a debate. *Bull. Amer. Meteor. Soc.*, 99, 837–839, <https://doi.org/10.1175/BAMS-D-17-0157.1>.
- Ralph, F.M., and Coauthors, 2020: West Coast forecast challenges and development of Atmospheric River Reconnaissance. *Bull. Amer. Meteor. Soc.*, 101, E1357–E1377, <https://doi.org/10.1175/BAMS-D-19-0183.1>.
- Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J.D. and McTaggart-Cowan, R., 2018. The North Atlantic Waveguide and Downstream Impact Experiment. *Bull. Amer. Meteor. Soc.*, 99, 1607–1637, <https://doi.org/10.1175/BAMS-D-17-0003.1>.
- Schindler, M., Weissmann, M., Schäfler, A. and Radnoti, G., 2020: The impact of dropsonde and extra radiosonde observations during NAWDEX in autumn 2016. *Mon. Wea. Rev.*, 148, 809–824, <https://doi.org/10.1175/MWR-D-19-0126.1>.
- Stone, R.E., Reynolds, C.A., Doyle, J.D., Langland, R.H., Baker, N.L., Lavers, D.A. and Ralph, F.M., 2020. Atmospheric River Reconnaissance Observation Impact in the Navy Global Forecast System. *Mon. Wea. Rev.*, 148(2), 763–782.
- Torn, R. D., and Hakim G.J., 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, 136, 663–677, <https://doi.org/10.1175/2007MWR2132.1>.
- Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 648 pp.
- Wu, X., Tallapragada, V., Lord, S., Wu, K., and Ralph, M., 2021. Impact of Atmospheric River Reconnaissance Dropsonde Data on NCEP GFS Forecast: A Case Study. *The 2021 Blue Book*. 1-23. http://bluebook.meteoinfo.ru/uploads/2021/sections/BB_21_S1.pdf.
- Yang, F., and Tallapragada, V., 2018. Evaluation of Retrospective and Real-time NGGPS FV3GFS Experiments for Q3FY18 Beta Implementation. AMS 25th Conference on Numerical Weather Prediction. Denver, Colorado, 5 June 2018, 5B.3. <https://ams.confex.com/ams/29WAF25NWP/webprogram/Paper345231.html>
- Yang, F., Tallapragada, V.S., Kleist, D.T., Chawla, A., Wang, J., Treadon, R. and Whitaker, J., 2021: On the Development and Evaluation of NWS Global Forecast Systems Version 16. In 101st American Meteorological Society Annual Meeting. AMS.
- Zheng, M., Monache, L.D., Wu, X., Ralph, F.M., Cornuelle, B., Tallapragada, V., Haase, J.S., Wilson, A.M., Mazloff, M., Subramanian, A.C., Cannon, F., 2020. Data Gaps within Atmospheric Rivers over the Northeastern Pacific. *Bull. Amer. Meteor. Soc.*, 102(3), E492–E524, <https://doi.org/10.1175/BAMS-D-19-0287.1>.
- Zheng, M., and Coauthors, 2022. Improved Forecast Skill through the Assimilation of Dropsonde Observations from the Atmospheric River Reconnaissance Program. *J. Geophys. Res. Atmos.*, accepted.
- Zhou, L., Lin, S.-J., Chen, J.-H., Harris, L. M., Chen, X., and Rees, S. L. 2019. Toward convective-scale prediction within the next generation global prediction system. *Bull. Amer. Meteor. Soc.*, 100(7), 1225–1243.